Protein-protein interactions in GCR1 signalling in *Arabidopsis thaliana*

Lihua Zhang

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

May 2008

To

My mother and father

Acknowledgements

I am extremely grateful to my supervisor Dr. Richard Hooley for his guidance, support, inspiration, encouragement and patience I received throughout this project. Without him, this project could not have been completed smoothly. I also greatly appreciate Dr. Baoxiu Qi for her contribution to this project, for teaching me so many useful techniques, and for her stimulating and instructive advices on this project and valuable comments on this thesis.

I deeply appreciate my second supervisor Professor Rod Scott and his family, for their kindest advice and support, and for taking care of me during my years in the UK. I would especially like to thank Dr. James Doughty, Dr. Paul Whitley, and Dr. Stefan Bagby for their help and suggestions on this project.

A big thank you to all my lab mates, both past and present, especially Dr. Ben Kemp, Dr. Pradeepika Saputhanthri, Dr. Marie Schruff, Dr. Melissa Spielman, Dr Sushma Tiwari, Lucille Brudenelle, Ahmed Bolbol, Ahlam Bouariky, Lindsay Dytham, Rhiannon Hughes and Ajaraporn Sriboonlert, for helping me and providing a pleasant working environment. My thanks also go to the project students who contributed to the work presented here: Richard Browning, John Locke, Callum Parr, Naomi Asantewa-Sechereh and Bicheng Yang.

I would like to express my deepest gratitude to my family and friends, especially my parents, for their love, encouragement and endless support.

Last but not least, I would like to acknowledge the Overseas Research Students Awards Scheme (ORSAS) and University of Bath for financial support.

Abstract

The G-protein coupled receptors (GPCRs) are seven-transmembrane receptors that transduce signals from the cell surface to intracellular effectors. There are more than 1000 GPCRs in metazoans, while no GPCR has been definitively identified in plants. The most promising plant GPCR candidate, *Arabidopsis* G-protein coupled receptor 1 (GCR1), physically couples to the G-protein α subunit GPA1 and is involved in cell cycle regulation, blue light and phytohormone responses, but its signalling network remains largely unknown. This project aimed to achieve a better understanding of GCR1 signalling by identifying its interactors using a novel yeast two hybrid system – the Ras Recruitment System (RRS). Screening of an *Arabidopsis* cDNA library using a bait comprising intracellular loop 1 (i1) and 2 (i2) of GCR1 resulted in the isolation of 20 potential interactors. Extensive reconfirmation screening demonstrated that three of these interactors: Thioredoxin *h*3 (TRX3), Thioredoxin *h*4 (TRX4) and a DHHC type zinc finger family protein (*zf*-DHHC1) interact specifically with both i1 and i2 of GCR1. This was supported by the reverse RRS (rRRS) and 6xHis-pull-down assays. It is speculated that TRX3 and TRX4, which can reduce disulfide bridges of target proteins and act as powerful antioxidants, may regulate GCR1-mediated signalling events in response to oxidative stress. Alternatively, they may modulate GCR1 targeting or signalling through their chaperone activities. *zf*-DHHC1 has a predicted membrane topography that is shared by most DHHC domain-containing palmitoyl acyl transferases. It may modify GCR1 activity through palmitoylation of the two cysteines located at the cytoplasmic end of the first transmembrane domain. Together, these findings contribute to the growing understanding of the GCR1 signalling network, and provide valuable starting points for further investigation.

List of Abbreviations

2-DE	2-dimensional gel electrophoresis
3D	3-dimensional
6xHis	hexahistidine
aa	amino acid
ABA	abscisic acid
AD	activating domain
AGB1	*Arabidopsis* G-protein β subunit
AGG1	*Arabidopsis* G-protein γ subunit 1
AGG2	*Arabidopsis* G-protein γ subunit 2
AGI	Arabidopsis Genome Initiative
amp	ampicillin
AP2	clathrin adaptor protein 2
APS	ammonium persulfate
APTs	protein palmitoyl thioesterases
ARF	ADP-ribosylation factor
ASK-1	apoptosis signal-regulating kinase-1
BCIP	5-bromo-4-chloro-3-indolyl phosphate
BD	binding domain
bp	base pair
BR	brassinosteroid
BSA	bovine serum albumin
cAMP	cyclic adenosine 3', 5'-monophosphate
CBP	calmodulin binding peptide
cDNA	complimentary DNA
Co-IP	Co-immunoprecipitation
Col-0	Columbia-0
Cter	C-terminus
ddH_2O	double distilled water
DIABLO	direct IAP binding protein with low pI
dicot	dicotyledonous species
DMSO	dimethyl sulfoxide

DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreitol
E.coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetic acid
EGTA	ethylene glycol tetraacetic acid
ER	endoplasmic reticulum
FTR	ferredoxin-dependent thioredoxin reductase
g	gravity
GA	gibberellic acid
GAPs	GTPase activating proteins
GCR1	G-protein coupled receptor 1
GCR2	G-protein coupled receptor 2
GDI	Guanine nucleotide Dissociation Inhibitor
GDP	guanosine 5'-diphosphate
GEFs	guanine nucleotide exchange factors
GFP	green fluorescent protein
GIPs	GPCR interacting proteins
GPA1	*Arabidopsis* G-protein α subunit
GPCRs	G-protein coupled receptors
GR	glucocorticoid receptor
GRASP	rapid affinity-capture of signalling proteins
GRKs	GPCR kinases
GST	glutathionine-S-transferase
GTP	guanosine 5'-triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMM	hidden Markov model
i1	intracellular loop 1
i2	intracellular loop 2
i3	intracellular loop 3
IAP	inhibitor of apoptosis protein
IP_3	inositol 1,4,5-triphosphate
IPTG	isopropyl-1-thio-β-D-galactoside

JAK	Janus tyrosine kinases
kan	kanamycin
kDa	kilodalton
LB	Luria-Bertani medium
MBP	maltose-binding protein
MIHA	inhibitor of apoptosis protein homolog A
min	minute(s)
MLO	mildew resistance locus O
monocot	monocotyledonous species
mRas	mammalian Ras
mRNA	messenger RNA
MS	mass spectrometry
NADPH	nicotinamide adenine dinucleotide phosphate
NBT	nitro blue terrazolium
NF-Y	nuclear factor Y
NMDAR	N-methyl-D-aspartate receptor
NMR	nuclear magnetic resonance
Nter	N-terminus
NTR	NADPH-dependent thioredoxin reductase
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PATs	palmitoyl acyl transferases
PCR	polymerase chain reaction
PD1	prephenate dehydratase 1
PDZ	PSD95-disc large-Zonula occludens
PEG	polyethylene glycol
PI-PLC	phosphatidylinositol-specific phospholipase C
PKA	protein kinase A
PKC	protein kinase C
RAMPs	receptor activity modifying proteins
RGS	regulator of G-protein signaling
RhoGDI	interactor of Rho guanine Nucleotide Dissociation Inhibitor
RNA	ribonucleic acid

ROS	reactive oxygen species
rpm	revolutions per minute
rRRS	reverse Ras Recruitment System
RRS	Ras Recruitment System
S1P	sphingosine-1-phosphate
SDS	sodium dodecyl sulfate
sec	second(s)
SNARE	soluble N-ethylmaleimide-sensitive fusion protein attachment
SPR	surface plasmon resonance
SRK	S Receptor Kinase
STAT	signal transducers and activators of transcription
TAP	tandom affinity purification
TE	Tris EDTA
TEMED	N,N,N',N'-Tetramethylethylenediamine
TEV	tobacco etch virus
TM	transmembrane
TNF-α	tumor necrosis factor-α
Tris	tris(hydroxymethyl)aminoethane
Tween 20	polyoxyethylene sorbitan monolaurate
U	unit
UV	ultra-violet
V	volts
v/v	volume to volume
Vol	volume
w/v	weight to volume
XLGs	extra-large G-proteins
YFP	yellow fluorescent protein
yRas	yeast Ras
β2AR	β2-adrenergic receptor

Contents

Chapter 1 General introduction

1.1 G-protein coupled receptors (GPCRs)

1.1.1 GPCR signalling in animals and other organisms

1.1.1.1 GPCRs in animals and other organisms

G-protein coupled receptors (GPCRs) are a super family of transmembrane proteins. They owe their name to their interaction with the heterotrimeric G-proteins (Marinissen and Gutkind, 2001). They belong to a more general class of proteins, the guanine nucleotide exchange factors (GEFs), which promote the GDP/GTP exchange on the G-protein (Assmann, 2002; Wieland and Michel, 2005; Ramachandran and Cerione, 2006). GPCRs have been identified in vertebrates, invertebrates, arthropods, insects, nematodes, fungi, yeast, viruses and plants (Hooley, 1998; Fredriksson and Schioth, 2005). They are the largest family of cell surface molecules involved in signal transduction, representing 1-5% of the invertebrate and vertebrate genomes. For example, GPCRs are encoded by 1% of total genes in *Drosophila melanogaster* (Adams *et al*., 2000) and 5% of all genes in *Caenorhabditis elegans* (Bargmann *et al*., 1998). More than 1% of the human genome encodes greater than 1000 GPCRs (George *et al*., 2002; Nambi and Aiyar, 2003; Nemoto and Toh, 2003).

A typical GPCR is composed of an extracellular N-terminus (Nter), seven transmembrane (TM) domains linked by three intracellular (i1, i2, and i3) and three extracellular (e1, e2, e3) loops, and a cytoplasmic C-terminus (Cter) (Figure 1.1). The most common feature for all GPCRs is the conserved 7TM structure (Chen *et al*., 2004; Fredriksson and Schioth, 2005). The first high resolution crystal structure of a mammalian GPCR, bovine rhodopsin, was solved in 2000 (Palczewski *et al*., 2000). It confirmed the predicted 7TM structure of rhodopsin and has been used as the general structural model for

all GPCRs (Ellis, 2004). The first high resolution structure of a human GPCR, the β2-adrenergic receptor (β2AR) was solved in 2007 (Cherezov *et al*., 2007; Rasmussen *et al*., 2007; Rosenbaum *et al*., 2007). The overall arrangement of the TM domains of the β2AR is similar to the rhodopsin, but the β2AR has a more open structure and differs from rhodopsin in having weaker interactions (an "ionic lock") between the cytoplasmic ends of TM3 and TM6, involving the conserved D(E)RY sequence (Rasmussen *et al*., 2007). The structural difference between these two proteins may reflect their difference in receptor activity (Rasmussen *et al*., 2007). At the same time, it reveals that although the rhodopsin-based model is very useful, it does not represent all GPCRs.

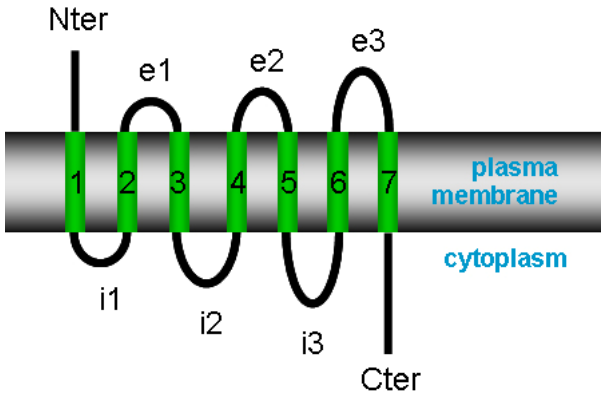


Figure 1.1 Schematic representation of a typical GPCR. A typical GPCR is composed of an extracellular N-terminus (Nter), seven transmembrane (7TM) domains linked by three intracellular (i1, i2, and i3) and three extracellular (e1, e2, e3) loops, and a cytoplasmic C-terminus (Cter).

In addition to the conserved 7TM domain, GPCRs have a number of other common features, such as the potential N-glycosylation sites in either or both of the N-terminal and e2 regions (Nakagawa *et al*., 2001; Lanctot *et al*., 2005). Glycosylation has been reported to play an important role in the cell surface expression, ligand binding, effector coupling, and quality control of many GPCRs (Walsh *et al*., 1998; Lanctot *et al*., 2005). The conserved cysteine residues that are located in e2 and e3, are thought to form a disulphide bridge, and are important for cell surface expression, ligand binding, receptor

activation and maintaining the secondary structure (Cook and Eidne, 1997; Zhang *et al*., 1999; Fredriksson, *et al*. 2003; Ray *et al*., 2004). Other conserved features include the LAXXD motif on TM2 which is thought to be involved in ligand binding and receptor cycling (Parker *et al*., 2008), the D(E)RY motif in TM3 that plays an important role in receptor activation and in the regulation of the receptor's interaction with its G-protein (Teller *et al*., 2001; Cotecchia *et al*., 2003), and the NPXXY in TM7 that contributes to G-protein coupled receptor internalization and signal transduction (He *et al*., 2001). Besides, there are also conserved serine and threonine residues in i3 or the C-terminal tail which are potential phosphorylation sites (Hauger *et al*., 2000; Miller *et al*., 2003; Charest and Bouvier, 2003; Potter *et al*., 2006; Torrecilla *et al*., 2007).

GPCRs can be classified into six families based on amino acid sequence similarity. These are Class A rhodopsin like, which account for over 80% of all GPCRs (Davies *et al*.,2007); Class B secretin like; Class C metabotropic glutamate/pheromone; Class D fungal pheromone; Class E cAMP receptors; and the Frizzled/Smoothened family (Horn *et al*., 1998). Some families contain a few subfamilies based on their ligand specificity, while others are small and not further subdivided (Josefsson, 1999; Qian *et al*., 2003). There is little sequence conservation across the six families, and sequence conservation within a family is limited to about 25% sequence identity within the TM domains (Bockaert *et al*., 2003; Jone and Assmann, 2004; Pandey and Assmann, 2004).

1.1.1.2 GPCR signalling and GPCR interacting proteins

GPCRs have a wide variety of ligands, including amino-acids, nucleotides, peptides, hormones, growth factors, odorants, Ca^{2+} (Bockaert and Pin, 1999; Marinissen and Gutkind, 2001; Bockaert *et al*., 2003). These ligands can be classified into three types: agonists, inverse agonists and antagonists (Prather, 2004; Okuno *et al*., 2006). Agonists bind to and activate GPCRs, as

well as stabilize the active state of GPCRs; inverse agonists bind to inactivated GPCRs and stabilize the inactive state of GPCRs; antagonists have equal preferences for both activated and inactivated GPCRs, and are able to block actions produced by either agonists or inverse agonists (Prather, 2004). The ligand binding sites include the extracellular and TM regions of GPCRs (Watson and Arkinstall, 1994).

Upon agonist binding, a GPCR adopts a conformational change in the TM domains that transmits to the intracellular regions where it interacts with the heterotrimeric G-protein to stimulate GDP/GTP exchange (Hammer *et al*., 2007). The heterotrimeric G-protein is composed of α, β and γ three subunits. The β and γ subunits are tightly associated and function as a Gβγ dimer. The Gα subunit is responsible for binding to a GPCR, but the association is greatly enhanced by Gβγ (Neer, 1995). The conformational change of the activated GPCR induces conformational change in its associated G-protein, which promotes the exchange of GDP for GTP in the Gα subunit (Neer, 1995; Kleanthous, 2000). The active, GTP-bound Gα dissociates from the GPCR and the Gβγ dimmer, so that Gα and Gβγ can each interact with their downstream effectors to alter cellular processes. The Gα subunit has an intrinsic GTPase activity that can hydrolyze GTP to GDP. The GDP-bound Gα becomes inactive and reassociates with the Gβγ dimer. The heterotrimer then returns to the membrane and reassociates with its receptor (Figure 1.2).

The G-proteins to which GPCRs couple are divided into two classes: the heterotrimeric and the small (monomeric) G-proteins (Ma, 1994). So far, 23 Gα, 6 Gβ and at least 12 Gγ subunits have been identified in humans (Vanderbeld and Kelly, 2000). The heterotrimeric G-proteins are classified into four categories (Gs, Gi/o, Gq/11, and G12) based on structural and functional similarities of the Gα subunit (Simon *et al*, 1991). By connecting receptors (e.g. GPCRs) and effectors (e.g. adenylyl cyclase), heterotrimeric G-proteins route extracellular signals to a network of intracellular signalling pathways that are involved in numerous biological processes, such as

regulation of metabolic enzymes, ion channels, transporters, etc. (Neves *et al*., 2002). In eukaryotes, more than 100 small G-proteins have been identified, which are structurally classified into at least five major families: Ras, Rho, Rab, Arf/Sar1 and Ran (Hall, 2000; Takai *et al*., 2001). Some of them, e.g. Ras and Rho can be activated by stimulated heterotrimeric G-proteins, while others e.g. Rab and Arf can directly interact with GPCRs (Bhattacharya *et al*., 2004). Small G-proteins have diverse functions, such as the regulation of intracellular signal transduction, cell proliferation and differentiation, cell polarity, cytoskeletal organization, the import and export of protein and RNA through the nuclear membrane (Hall, 2000; Pandit and Srinivasan, 2003).

In addition to G-proteins, GPCRs interact with many other proteins that either modify GPCR activities or mediate G-protein independent signalling pathways. For instance, activated GPCRs can be rapidly phosphorylated by GPCR kinases (GRKs) or second messenger-activated kinases, for instance, protein kinase A (PKA) and protein kinase C (PKC) (Figure 1.2; Nakagawa *et al*., 2001; Thomas and Qian, 2003). The β-arrestin proteins can bind to phosphorylated GPCR, promoting the uncoupling of the receptor from the G-protein which results in attenuated sensitivity to ligand, a phenomenon termed GPCR desensitization (Ferguson, 2001; Gainetdinov *et al*., 2004). The binding of β-arrestin also promotes the removal of the receptor from the membrane by clathrin-mediated endocytosis (Charest and Bouvier, 2003; Suga and Haga, 2007). The internalized GPCRs are either degraded in lysosome or recycled back to cell surface (Figure 1.2; Seethala and Fernandes, 2001; Bernard *et al*., 2006). Other GPCR interacting proteins (GIPs) include the receptor activity modifying proteins (RAMPs), the PDZ (PSD95-disc large-Zonula occludens) domain containing proteins, ionic channels, and ionotropic receptors (Brady and Limbird, 2002; Bockaert *et al*., 2003; Bockaert *et al*., 2004a; Bockaert *et al*., 2004b). A growing body of evidence indicates that a GPCR can also interact with itself (homodimerization) or with another GPCR (heterodimerization), and this

process is suggested to be important for receptor maturation and internalization (Terrillon and Bouvier, 2004).

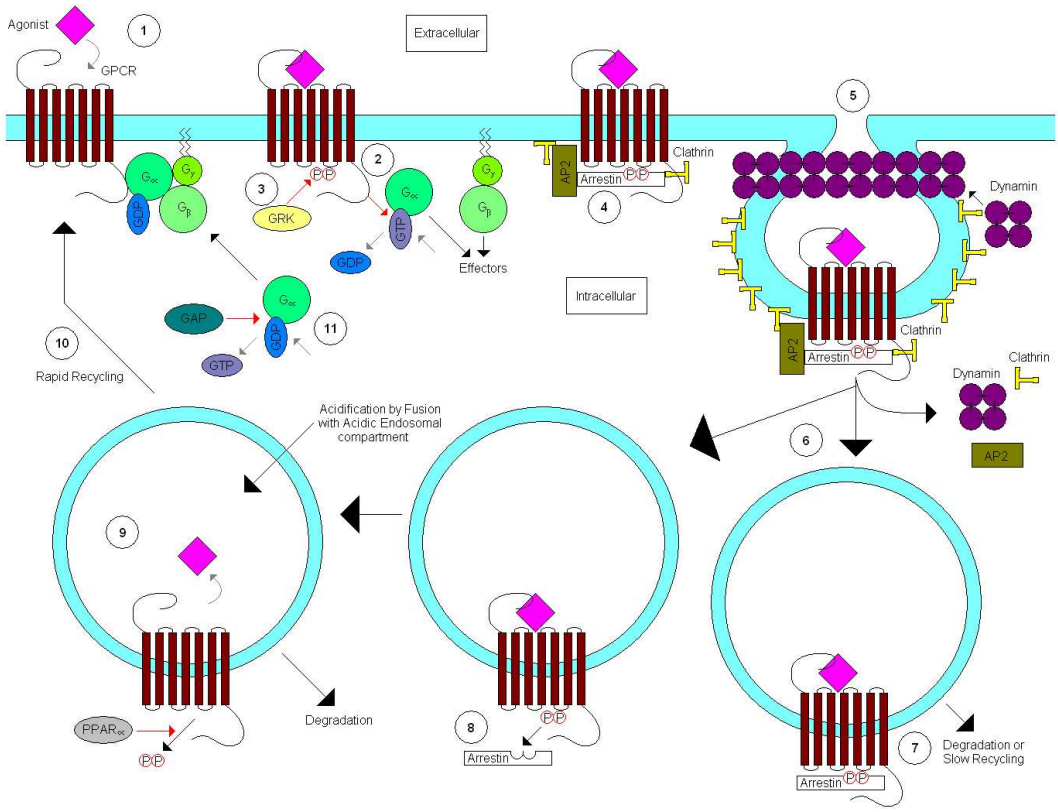


Figure 1.2 GPCR trafficking in animals (by project student Richard Browning, 2007). Ligand binding activates GPCR (1), which in turn promotes the GDP/GTP exchange in Gα and subsequent disassociation of Gα from GPCR and Gβγ (2). Arrestin binds to GRKs (or second messenger-activated kinases) phosphorylated GPCR (3) to prevent further G-protein activation. Arrestin can also bind the clathrin adaptor protein AP2, and clathrin (4) causing a gradual invagination of the receptor in a clathrin-coated pit (5), eventually leading to dynamin-mediated endocytosis. This can then follow one of two pathways (6) depending upon GPCR type, both accompanied by loss of dynamin molecules, AP2 and clathrin. If the GPCR-arrestin complex is stable (7), the vesicle will be sequested and the GPCR will be degraded. If the GPCR-arrestin complex is unstable (8), the arrestin will rapidly disassociate upon internalisation. The vesicle is then acidified by fusion with an acidic endosomal vesicle compartment, which dissociates the ligand, while the phosphorylase PPARα removes the phosphates (9). The vesicle is then rapid recycled back to the membrane (10). The G-protein is recycled (11) by hydrolysis of the GTP by inherent Gα GTPase activity which may be enhanced by GTPase activating proteins (GAPs). This allows the reassociation of Gα with Gβγ, and subsequent association with GPCR.

GIPs are known to interact with the intracellular loops, transmembrane domains and the C-terminal tails of GPCRs. For example, the second and third intracellular loops of the α-2A-adrenergic receptor are responsible for its interaction with GRK2 (Pao and Benovic, 2005). The transmembrane domains represent the major dimerization interface for both GPCR homodimers and heterodimers (Bulenger *et al*., 2005). The C-terminal tail is recognised as the main domain responsible for interaction with GIPs, and more than 50 GIPs have been identified to interact with this “magic tail” (Bockaert *et al*., 2004b).

GPCRs are involved in a wide variety of biological processes in animals, such as neurotransmission, visual perception, smell, taste, embryogenesis, development, secretion, metabolism, and immune and inflammatory responses (Kleanthous, 2000; Wong, 2003). Mutation in GPCRs causes many diseases in humans (reviewed in Schoneberg *et al*., 2004; Insel *et al*., 2007). Their large numbers and important roles have drawn attention from not only scientists but also drug developers. GPCRs are the targets of more than 50% of the current therapeutic agents, representing more than $50 billion in annual global sales (Marinissen and Gutkind, 2001; Lundstrom, 2006).

1.1.2 GPCR signalling in plants

1.1.2.1 GPCRs in plants

Considering the large number of GPCRs in animals, one might be interested in how many GPCRs there are in plants. To date, no GPCR has been definitively identified in plants (Chen *et al*., 2004). The most promising GPCR candidate in *Arabidopsis thaliana*, GCR1, was identified independently by two research groups using sequence-based homology searches of the EST database with the representatives of the six GPCR families as queries (Josefsson and Rask, 1997; Plakidou-Dymock *et al*., 1998). The predicted overall topology of GCR1 is similar to GPCRs that contain a 7TM domain with

a preferred orientation of an extracellular Nter and an intracellular Cter. GCR1 has 18-23% amino acid identity (46-53% similarity) to the *Dictyostelium discoideum* cAMP receptors (Plakidou-Dymock *et al*., 1998). It also has significant amino acid sequence similarity in certain regions to Class B receptors for corticotrophin releasing factor and calcitonin, to rhodopsin, serotonin and olfactory receptors of Class A, to members of the Frizzled/Smoothened family, and to Methuselah-like proteins of *Drosophila* (Josefsson and Rask, 1997; Plakidou-Dymock *et al*., 1998; Pandey and Assmann, 2004).

GCR1 has a number of amino acids that are conserved in many GPCRs, as summarised by Josefsson and Rask (1997) and Plakidou-Dymock (1998). It has a potential N-glycosylation site in e2, which is consistent with many GPCRs that have potential N-glycosylation sites in either or both of the N-terminal domain and e2 (Nakagawa *et al*., 2001; Lanctot *et al*., 2005). There are conserved cysteines located on e2 and e3 in the majority of GPCRs (Cook and Eidne, 1997; Zhang *et al*., 1999; Fredriksson, *et al*. 2003; Ray *et al*., 2004). Similarly to these GPCRs, there are cysteines in e1 and e2 of GCR1, which could form a disulphide bridge. Other conserved features include the LAXXD motif in TM2; an arginine at the boundary between TM3 and i2, a tryptophan in TM4, the motif FXXPXXXXXXY in TM5, asparagine and tyrosine in TM7, and multiple serines and a threonine in the C-terminal tail, all of which are conserved in many GPCRs (Colson *et al*.,1998; Hauger *et al*., 2000; Miller *et al*., 2003; Charest and Bouvier, 2003; Potter *et al*., 2006; Torrecilla *et al*., 2007; Graaf *et al*., 2008; Parker *et al*., 2008).

GCR1 is a single copy gene AT1g48270 (the genomic DNA is 2468 bp) (Plakidou-Dymock *et al*., 1998). It is expressed at a very low level in roots, stem, leaves, guard cell protoplasts and mesophyll cell protoplasts (Plakidou-Dymock *et al*., 1998; Colucci *et al*., 2002; Pandey and Assmann, 2004), whereas its highest expression is found in meristematic regions, e.g. flowering buds and small siliques of 5-week old plant (Colucci *et al*., 2002).

The subcellular localization of GCR1, or more specifically, the association of GCR1 to the plasma membrane was first revealed by detecting GCR1-GFP (green fluorescent protein) fluorescence to the outer edge of the leaf epidermal cells of *Arabidopsis* plants using confocal laser-scanning microscopy (Humphrey and Botella, 2001). Using a similar method, Chen *et al*. (2004) detected GCR1-GFP fluorescence in a punctate pattern around the plasma membrane, indicating its membrane association and possible internalization (Chen *et al*., 2004). The physical association with G-protein is another important factor for the classification of GCR1 as a bona fide GPCR (Humphrey and Botella, 2001). To date, one Gα-subunit (GPA1), one Gβ-subunit (AGB1) and two Gγ-subunits (AGG1 and AGG2) of the heterotrimeric G-protein have been identified in *Arabidopsis* (Jones and Assmann, 2004). Pandey and co-workers (2004) demonstrated by *in vitro* pull-down, split-ubiquitin yeast two-hybrid, and co-immunoprecipitation that GCR1 and GPA1 are physically coupled, and their interaction depends on the intracellular domains of GCR1. Altogether, the 7TM topography and conserved feature that are common to known GPCRs, the association with plasma membrane, and the physical interaction with GPA1, all define GCR1 as a very promising plant GPCR candidate. However, to date there has been no demonstration of GCR1 interaction with GPA1 causing a GDP/GTP exchange reaction which is essential for demonstrating a role of GCR1 as a GPCR.

In addition to GCR1, there are many other 7TM proteins in plants. It has been predicted that there are approximately 6,500 putative integral membrane proteins among the 25,498 *Arabidopsis* proteins (Schwacke *et al*, 2003). Using multiple nonalignment approaches, Moriyama and co-workers (2006) identified 394 *Arabidopsis* proteins as 7TM receptors, 54 of which have an extracellular Nter. Of the list of these 54 proteins, GCR1 is the only one that shares similarities to known GPCRs (Temple and Jones, 2007). Another protein from the list named RGS1, which is characterized as the sole regulator of G-protein signaling (RGS) (Chen *et al*., 2003), has a predicted structure similar to a GPCR as well as an RGS box with GTPase accelerating

activity, and it interacts with the heterotrimeric G-Protein both *in vitro* and *in vivo* (Chen *et al*., 2003). It thus has been proposed to be a candidate GPCR (Assmann, 2005). RGS1 acts through GPA1 to modulate cell division in the primary root (Chen *et al*., 2006a), functions in GPA1 regulated sugar sensing pathways (Chen and Jones, 2004; Johnston *et al*., 2007b), and enhances abscisic acid (ABA) mediated root elongation, drought tolerance and seed germination in *Arabidopsis* (Chen *et al*., 2006b, c). Another seven proteins from the list of 54 7TM receptors belong to the 'mildew resistance locus O' (MLO) family (Moriyama *et al*., 2006). The MLO family is the only abundant family of 7TM proteins unique to plants, and a barley MLO is the only plant protein that has been experimentally confirmed to consist of a 7TM domain (Devoto *et al*., 1999). There are 15 MLO proteins in *Arabidopsis*, but none shares significant sequence similarity with GCR1 or any known GPCRs (Devoto, *et al*., 1999; 2003). They are proposed to be candidate GPCRs based on their membrane topology, subcellular location and domain-specific sequence variability (Devoto *et al*., 1999). MLOs have been shown to interact with calmodulin to regulate defence against mildew in barley in a G-protein independent manner (Kim *et al*., 2002a, b). Whether any MLO couples to G-proteins is still unclear (Assmann, 2005).

Recently, two GPCR candidates have been reported, one in *Araibdopsis* (Liu *et al*., 2007a) and one in Pea (*Pisum sativum*) (Misra *et al*., 2007). The *Arabidopsis* protein, named GCR2, was reported to be a putative GPCR, as it has predicted 7TM domain and interacts with GPA1 to mediate all known ABA responses in *Arabidopsis* (Liu *et al*., 2007b). However, based on in silico modelling and bioinformatics prediction, Johnston *et al*. (2007a) argued that although GCR2 might be both an ABA receptor and a G-protein modulator, it is neither a transmembrane protein nor a GPCR, but an *Arabidopsis* homolog of bacterial lanthionine synthetases. In response to this comment, Liu *et al*. (2007a) demonstrated that 9 out 12 computational programs predicted GCR1 to be a 7TM protein, that GCR2-YFP fusion protein is only localized in the plasma membrane, and that GCR1 interact with both GPA1 and AGB1 by

split-ubiquitin assay in yeast, all of which support the role of GCR2 as a putative GPCR in *Arabidopsis*. They also stated that although they can not rule out the possibility that GCR2 is a lanthionine synthetase, their data indicate that GCR2 may define a new type of GPCR (Liu *et al*., 2007a). The putative pea GPCR (PsGPCR) also has predicted 7TM domain, and it has 50% amino acid sequence identity with GCR1, mostly in the transmembrane regions (Misra *et al*., 2007). In addition, PsGPCR interacts with all the subunits of pea G-proteins (PsGα1, PsGβ and PsGγ1) and with itself, indicating that it might be a putative GPCR in pea (Misra *et al*., 2007). Apparently, a great deal of analysis needs to be performed to assess whether GCR2 and PsGPCR are bona fide plant GPCRs.

1.1.2.2 GCR1 signalling and GCR1 interacting proteins in *Arabidopsis*

Metazoan GPCRs can be activated by a large number of ligands. However, to date, no ligand of GCR1 has been identified (Apone *et al*., 2003; Jones and Assmann, 2004). Nevertheless, in the past few years, promising progress has been made towards understanding GCR1 signalling in *Arabidopsis*, which has reinforced its role as the most promising GPCR candidate in plants.

The phytohormone ABA is involved in many physiological processes, such as promotion of stomatal closure and inhibition of stomatal opening, promotion of seed dormancy and inhibition of seed germination (reviewed in Rock 2000; Finkelstein *et al*., 2002). Pandey and co-workers (2004) showed that GCR1 mutant *gcr1* is drought tolerant and is hypersensitive to plant hormone ABA and lipid metabolite sphingosine-1-phosphate (S1P, a ligand for GPCR in mammals) regulation of stomatal responses. Based on the observation that *gcr1* exhibits hypersensitivity, whereas *gpa1* exhibit insensitivity in the aspect of ABA and S1P responses in guard cells, they hypothesized that GCR1 might be a negative regulator of GPA1 in guard cells, or that GCR1 negatively regulates ABA signalling via a mechanism independent of its binding to GPA1 (Pandey and Assmann, 2004). They then studied the roles of GCR1, GPA1

and AGB1 in ABA control of seed germination and early post-germination growth and development. They found that *agb1* have greater ABA hypersensitivity than *gpa1* mutants, indicating that AGB1 might be the predominant regulator of ABA signaling and that GPA1 affects the action of AGB1 (Pandey *et al*., 2006). They also observed that *gcr1 gpa1* double mutants exhibit a *gpa1* phenotype and *agb1 gcr1* and *agb1 gcr1 gpa1* mutants exhibit an *agb1* phenotype, which indicates that GCR1 acts upstream and in concert with GPA1 and AGB1 for ABA signaling pathways during germination and early seedling development (Pandey *et al*., 2006). Above information implies that plants may have cell and tissue specific GCR1 signalling pathways in response to certain signals, such as ABA (Pandey *et al*., 2006).

By comparing the expression of *GCR1* to that of cyclin D2 (*CYCD2*), a cell cycle associated regulator of protein kinase activity that is expressed in early G1, Colucci and colleagues (2002) found that *GCR1* has a similar expression pattern as *CYCD2*, which suggests that *GCR1* has a cell cycle associated expression pattern. In addition, overexpression of GCR1 in tobacco Bright Yellow 2 (BY-2) cells increases the incorporation of thymidine into DNA and elevates the mitotic index of the cells, whereas overexpression of GCR1 in *Arabidopsis* creates transgenic plants that have shortened flowering time and produce seeds that lack dormancy (Colucci *et al*., 2002). Altogether, these results imply that GCR1 plays a role in the regulation of cell cycle and that the phenotype of GCR1 transgenic plants is the result of a modulation of the cell cycle (Colucci *et al*., 2002). Further investigation demonstrated that phosphatidylinositol-specific phospholipase C (PI-PLC) activity and inositol 1,4,5-triphosphate (IP_3) levels are higher in GCR1-overexpressing cells, that IP_3 levels are regulated by GCR1 during the cycle, and that PI-PLC inhibitor affects the rate of DNA synthesis, which indicates that GCR1 enhanced thymidine incorporation into DNA (enhanced DNA synthesis) involves PI-PLC as an effector and IP_3 as a second messenger (Figure 1.3; Apone *et al*., 2003). In addition, PI-PLC activity and IP_3 levels are higher in BY2 cells

overexpressing GPA1 than those overexpressing GCR1, suggesting that both GCR1 and GPA1 can affect the same signalling pathway that leads to DNA synthesis and entry into the cell cycle through PI-PLC activation (Apone *et al*., 2003). However, since the physical interaction between GCR1 or G-protein and PLC awaits assessment, it is also possible that GCR1 activates PI-PLC in an indirect way (Apone *et al*., 2003).

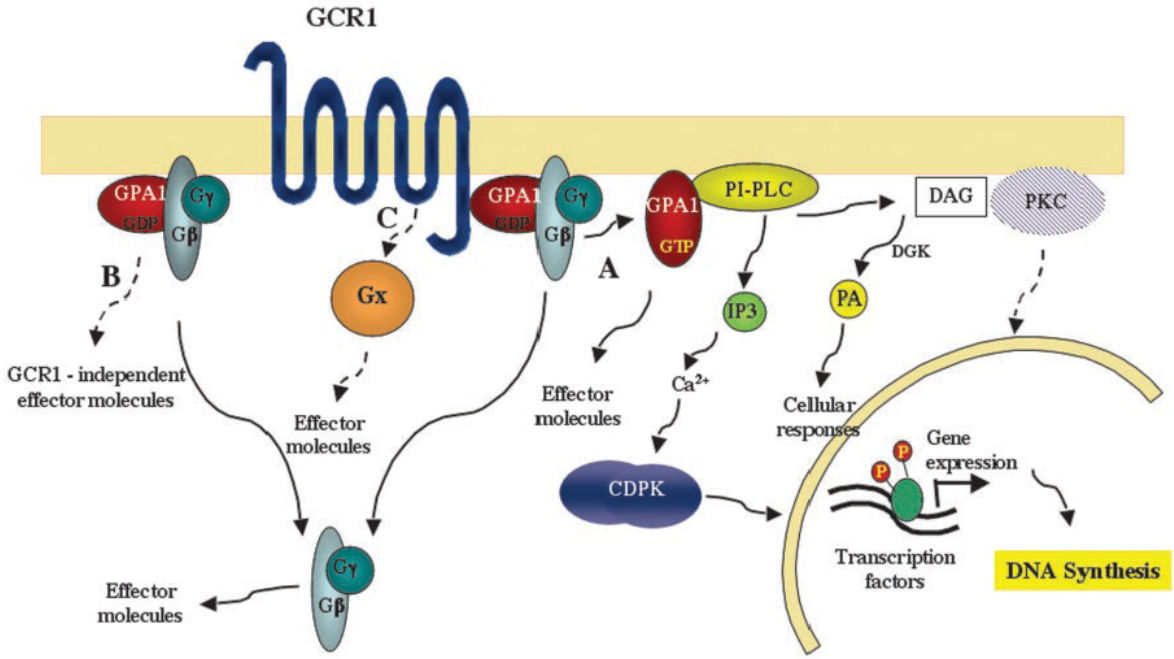


Figure 1.3 Speculative model showing how GCR1 signalling may lead to DNA synthesis (taken from Apone *et al*, 2003).

Light is one of the most important environmental factors that regulate numerous aspects of plants development, such as seed germination and the onset of flowering (Christie, 2006). Kaufman's lab identified that GCR1 is involved in blue light mediated events. In *Arabidopsis*, a single pulse of low-fluence blue light ($<10^{-1}$ μmolm^{-2}, equivalent in intensity to 1s of full moonlight) enhances the accumulation of phenylalanine (Phe), which is required for protein synthesis and serves as the precursor to a large number of secondary metabolites (Warpeha *et al*., 2006). Prephenate dehydratase 1 (PD1, also termed ADT3 by Cho *et al*., 2007) is an enzyme that presents in etiolated

seedlings and catalyzes the conversion of prephenate to phenylpyruvate, which is the immediate precursor to Phe (Warpeha *et al*., 2006). It has been shown that activated GPA1 strongly interacts with PD1 and increases its activity, and that both GCR1 and GPA1 participate in blue light mediated synthesis of Phe, because neither *gcr1* nor *gpa1* mutants accumulate Phe (Warpeha *et al*., 2006). Based on the above observations and the known interaction between GCR1 and GPA1, Warpeha *et al*. (2006) proposed that GCR1, GPA1 and PD1 may form all of or part of a signalling pathway that is responsible for blue light mediated synthesis of Phe and Phe-derived metabolites. They also demonstrated that ABA can act through the same pathway in etiolated cotyledons of *Arabidopsis* (Warpeha *et al*., 2006).

It was recently reported by Kaufman's group that a signal transduction chain, which consists of GCR1, GPA1, PRN1 and NF-Y, mediates both blue light and ABA responses in *Arabidopsis* (Warpeha *et al*., 2007). PRN1 (Pirin1), which interacts with GPA1 (Lapik and Kaufman, 2003), was originally identified through its interaction with NF-Y (Wendler *et al*., 1997). NF-Y (nuclear factor Y, also termed CBF) is a CCAAT box binding protein that consists of NF-Y-A, -B and -C three subunits (Mantovani, 1998; Silvio *et al*., 1999). A single pulse of low-fluence blue light induces the expression of several nuclear-coded genes, including the *light-harvesting chlorophyll a/b-binding protein* (*Lhcb*) gene family (Warpeha *et al*., 2007). Both the NF-Y-B9 subunit which physically interacts with PRN1, and the NF-Y-A5 subunit are required for blue light induced *Lhcb* expression (Warpeha *et al*., 2007). Based on these observations, Warpeha *et al*. (2007) proposed a model for the GCR1, GPA1, PRN1, NF-Y signal transduction mechanism that is responsible for blue light induced *Lhcb* expression (Figure 1.4). In this model, blue light is passed through the plasma membrane via GCR1 and GPA1, which subsequently activates PRN1, a shuttle protein that is capable of entering the nucleus to interact with NF-Y-B9 and NF-Y-A5 and a particular NF-Y-C subunit to facilitate *Lhcb* transcription. In addition, ABA can elicit *Lhcb* expression in etiolated seedlings through the same signalling pathway, and

this pathway may also inhibit ABA-mediated delay in seed germination (Warpeha *et al*., 2007).

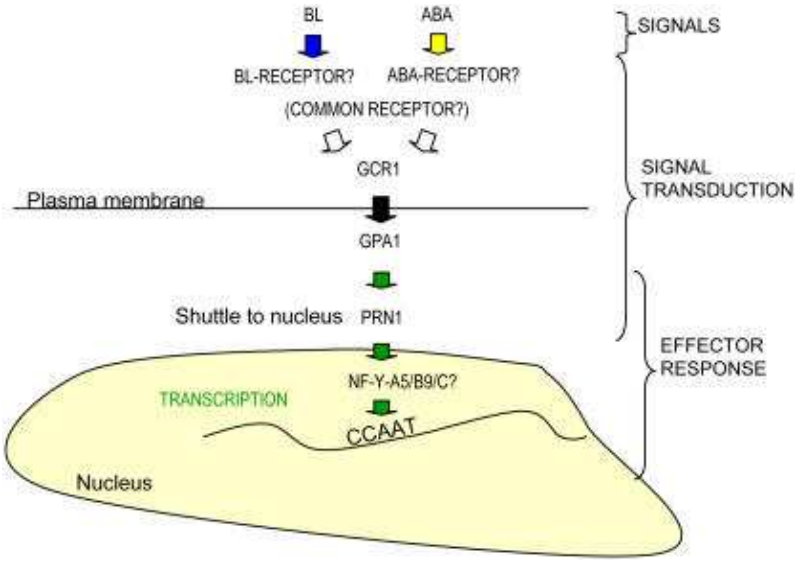


Figure 1.4 A proposed model for the GCR1, GPA1, PRN1, NF-Y signal transduction mechanism responsible for blue light (BL) induced of *Lhcb* transcription (taken from Warpeha *et al*., 2007).

In metazoa, GPCRs mediate biological responses that are independent of G-proteins. This type of GPCR signalling pathway also exists in *Arabidopsis*. It is known that the phytohormones gibberellic acid (GA) and brassinosteroid (BR) promote seed germination (reviewed in Kucera *et al*., 2005). Based on the evidence that GCR1 is coupled to GPA1 (Pandey and Assmann, 2004) and that both *gpa1* (Ullah *et al*., 2001) and *gcr1* (Chen *et al*., 2004) have reduced sensitivity to GA in seed germination, one would expect the *gcr1 gpa1* double mutants to display an epistatic phenotype. However, *gcr1 gpa1* double mutants, *agb1 gcr1* double mutants and *agb1 gcr1 gpa1* triple mutants have mostly synergistic effects of reduced sensitivity to GA and BR in seed germination, implying that GCR1 acts independently of G-protein in response to BR and GA in *Arabidopsis* seed germination (Chen *et al*., 2004).

As stated earlier, one Gα, one Gβ and two Gγ subunits of the heterotrimeric G-protein have been identified in *Arabidopsis*. In fact, the heterotrimeric G-protein subunits have also been found in other plants, including several dicots, e.g. tomato, lotus, pea, soybean and tobacco, lupin and spinach, and monocots e.g. rice, wild oat (reviewed in Assmann, 2002). The plant heterotrimeric G-proteins regulate cell proliferation and ion channels, and are involved in plant responses to hormones (e.g. GA, auxin, ABA, BR, ethylene), light, pathogens and drought, as well as some developmental events, such as seed germination, lateral root formation, hypocotyls elongation, hook opening, leaf expansion and silique development (reviewed in Jones, 2002; Iwasaki *et al*., 2003; Jones and Assmann, 2004; Perfus-Barceoch *et al*., 2004). In addition to the canonical heterotrimeric G-protein, seven noncanonical, extra-large G-proteins (XLGs) have been identified in plants, three in *Arabidopsis* and four in rice (Temple and Jones, 2007). They share significant amino acid sequence similarity over the C-terminal 633 residues, and the sequence identity within the Gα domain among these XLGs is 47% on average (Temple and Jones, 2007). Another group of plant proteins, the receptor for activated C-kinase 1 (RACK1s or Arcs), has predicted three dimensional structures similar to that of Gβs (Ishida *et al*., 1993; Assmann, 2005). But no XLG or RACK has been shown to interact with the canonical G-protein subunits, therefore there role as noncanonical G-proteins are hypothetical (Assmann, 2005), and their interaction with putative GPCRs awaits exploration.

Beside the G-proteins, no other GCR1 interacting proteins have been identified so far. The presence of potential phosphorylation sites in GCR1 suggest that it may be phosphorylated by GRKs or second messenger-regulated kinases, such as PKA and PKC, followed by arrestin mediated desensitization. However, no GRKs, PKA and PKC homologues, and no arrestin-like proteins have been identified in plants (Andreeva and Kutuzov, 2001; Gurevich and Gurevich, 2006). GCR1 might be regulated by kinases and arrestins that are structurally distinct from those in animals (Andreeva and Kutuzov, 2001).

Compared to our comprehensive understanding about GPCRs in animals, our knowledge about GCR1 in plants is limited. Generally, interacting proteins can be expected to be involved in the same cellular process (Coates and Hall, 2003). On this basis, an important aspect in gaining insight into the function of GCR1 is the identification of other proteins with which it interacts. The identification of these interactors could provide significant clues to the cellular pathways in which GCR1 participates.

1.2 Approaches for studying protein-protein interactions

Over the past several decades, numerous techniques have been developed for protein-protein interaction analysis, from biochemical approaches such as affinity chromatography and co-immunoprecipitation, to molecular genetic approaches such as the yeast two-hybrid system. There is no need to cover all the approaches here, therefore, some of the most commonly used large scale and high-throughput methods are addressed in the following sections.

1.2.1 Affinity chromatography

Affinity chromatography is designed to separate out all the proteins with the same specificity from a mixture. For example, actin affinity chromatography has been used in plant research to identify novel actin-binding proteins and confirm previous findings obtained with other techniques (Vitale, 2002). The rationale of affinity chromatography is illustrated in figure 1.5. The protein mixture, such as cellular or tissue extract, is passed through a column containing beads (eg. agarose, or sephadex) which have covalently attached ligand. Proteins that do not bind to the ligand will be washed through, while proteins that bind to the ligand will be retained. When all the unbound proteins have been washed through, the bound protein is eluted by adding a solution containing free ligand. This competes with the solid phase-bound substrate for the binding sites on the protein. The protein bound to the free ligand will be washed through, and the elution will be further analysed.

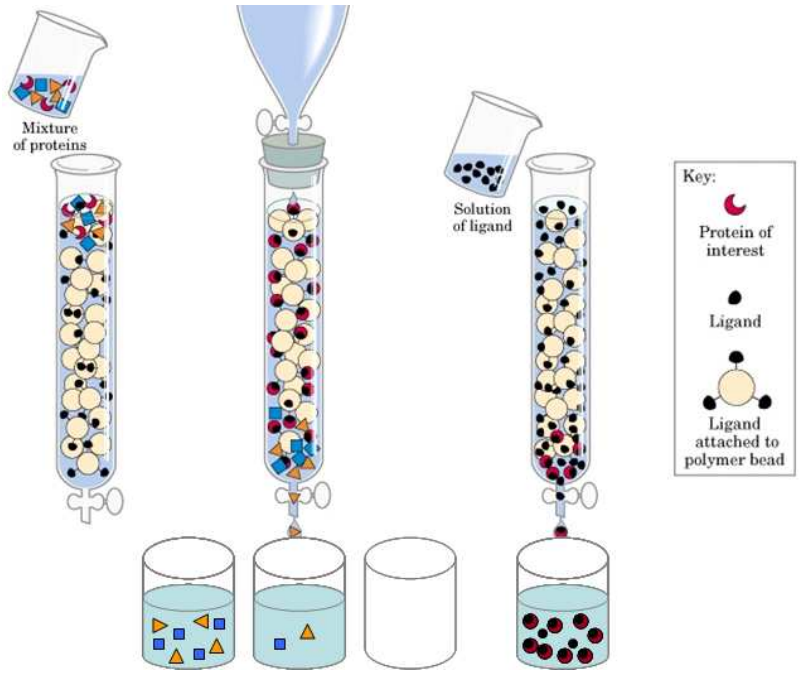


Figure 1.5 Rationale of affinity chromatography (adapted from Robertus, 2008). The protein mixture is passed through a column containing beads that have covalently attached ligand. Proteins that do not bind to the ligand will be eluted, while proteins that bind to the ligand will be retained. When all the unbound proteins have been washed through, the bound protein is eluted by adding a solution containing free ligand. The protein bound to the free ligand will be washed through, and the elution will be further analysed.

Affinity chromatography has been adapted to retrieve proteins associated with a tagged bait protein and subsequently to identify the individual components by mass spectrometry (MS), sequencing or Western blot. Such approaches have been applied for large scale protein-protein interaction analysis in prokaryotes and eukaryotes (O'Connor and Hames, 2008). Two of the mostly widely used methods, pull-down and tandem affinity purification (TAP), are described in detail below.

1.2.1.1 Pull-down

In a pull-down assay, the protein of interest (bait) is fused to a tag which allows it to immobilize to suitable affinity beads. The immobilized bait protein is incubated with either whole-cell or fractionated cell lysate to allow protein-protein interactions to occur (Lopper and Keck, 2007; figure 1.6). The

complex containing the bait protein and any associated proteins is washed to remove none specifically bound proteins, and subsequently removed from the affinity support using SDS-PAGE sample buffer or a competitive compound specific for the tag on the bait protein (Figure 1.6). The protein complex can be analysed by MS, sequencing or Western blot, depending on the purpose of the experiment.

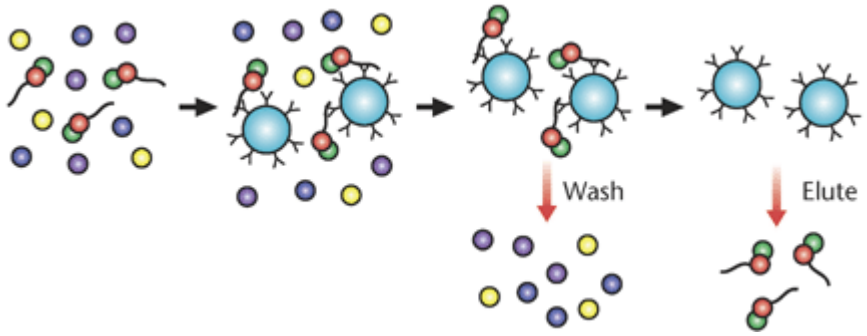


Figure1.6 The rationale of a pull-down assay (taken from Lopper and Keck, 2007). A pull-down assay allows selective capture of the tagged protein of interested (red circles) and its binding partners (green circles) onto an affinity support.

The most commonly used tags include polyhistidine (His), glutathione-*S*-transferase (GST), maltose-binding protein (MBP) and a number of epitopes that can be recognized by specific antibodies, e.g. Myc and FLAG (O'Connor and Hames, 2008). Of these, the His-tag remains a popular choice for a number of reasons. The His-tagged fusion protein contains a polyhistidine stretch that can bind to divalent metal ions such as nickel and cobalt with high affinity (Hochuli *et al*., 1987). The small size of the His-tag (normally 4-6 amino acids) reduces the chance to alter the structure or activity of a protein of interest, and the chance of non-specific binding. The fusion protein can be denatured during the purification and the His-tag will remain competent to bind to metal ions, therefore, the His-tag system has a distinct advantage for analysing insoluble proteins (Golemis and Adams, 2005).

Pull-down has been widely used as an initial screening assay for identifying novel protein-protein interactions. For example, using large scale pull-down

screen together with mass spectrometry analysis, Fang *et al*. (2004) identified mouse Sur2 which is a subunit of the Mediator complex, as a binding partner to mouse adenovirus type 1 early region 1A, which encodes a virulence gene in viral infection of mice. In another study, a large-scale comprehensive pull-down assay was performed using a His-tagged *E.coli* ORF clone library. Of the 4339 bait proteins tested, 11511 interacting partners were identified for 2667 proteins that were successfully overproduced and purified (Arifuzzaman *et al*., 2006). In addition, pull-down has also been used extensively as a confirmation tool to evaluate known or suspected protein-protein interactions.

Although pull-down is a valuable approach for studying protein-protein interactions, there are a number of points to be considered when carrying out an experiment. Firstly, it may detect false interactions caused by relatively high concentrations of proteins or by unsuitable binding and washing conditions used in the experiment (Picard, 1999). Secondly, interactions detected by *in vitro* pull-down experiments may not exist *in vivo*, simply because these interacting proteins never encounter one another in the cell. Lastly, it may fail to detect certain interactions due to the absence of cofactors, unsuitable experiment conditions, or due to protein misfolding when overexpressed in *E.coli* or when synthesized *in vitro*.

1.2.1.2 Tandem Affinity Purification (TAP)-tagging

The TAP-tagging method relies on two affinity purification steps that are applied sequentially. It was developed for high yield purification of protein complexes formed under native conditions (Abe *et al*., 2008; O'Connor and Hames, 2008). The TAP tag is composed of two affinity components, a calmodulin binding peptide (CBP) and two immunoglobulin (IgG)-binding domains of protein A, separated by a protease cleavage site derived from tobacco etch virus (TEV). Extracts containing the tagged protein which are prepared in a natural host cell or organisms are used in a TAP-tagging experiment. In the first purification step, the tagged protein and its associated

partners are purified from the extract by the binding of protein A to the IgG beads (Azarkan *et al*., 2007; figure1.7). The bound proteins are subsequently released from the IgG beads via TEV protease mediated cleavage. In the second purification step, CBP on the tagged protein binds to the calmodulin beads in the presence of calcium ions. Finally, the tagged protein and its associated partners are eluted from the calmodulin beads by calcium withdrawal (using ethylene glycol tetraacetic acid, EGTA) or by boiling in SDS sample buffer (O'Connor and Hames, 2008; figure1.7). The components of the protein complex can be resolved by SDS-PAGE and identified by MS.

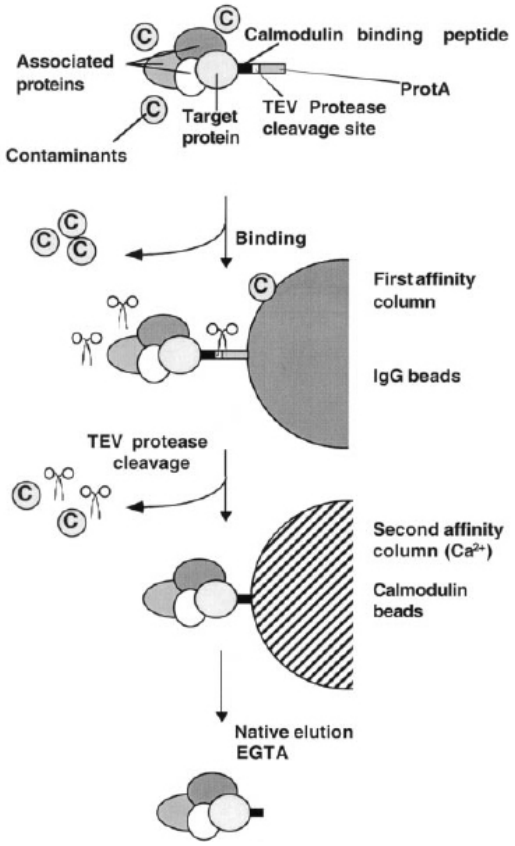


Figure 1.7 Schematic representation of the TAP purification steps (taken from Azarkan *et al*., 2007).

TAP-tagging has been used successfully in *S.cerevisiae*, *E.coli*, *C.elegans*, *Drosophila*, mouse, human and plants (Rohila *et al*., 2006; O'Connor and Hames, 2008). It was first applied to plants by Rohile *et al*. (2004) who

identified HSP90 and HSP70 as the interacting proteins of the glucocorticoid receptor in tobacco. Using an alternative TAP-tagging strategy, where the TEV protease cleavage site was replaced with a low temperature active rhinovirus 3C protease site and the CBP domain was replaced with two affinity tags (6x-His and 9x-Myc), Rubio *et al*. (2005) demonstrated that the components of the multi-protein COP9 signalosome (CSN) complex can be successfully co-purified from *Arabidopsis* extracts with the CSN3 subunit. In a large scale application, partners were found for 23 (out of the 39) TAP-tagged rice kinases, and some of these interactions were consistent with known protein complexes found in other species (Rohila *et al*., 2006). The increasing number of examples indicates that TAP-tagging is an effective method for detecting novel protein-protein interactions in plants (Abe *et al*., 2008).

One of the distinct advantages of using TAP-tagging to study protein-protein interactions is that it allows rapid purification of protein complexes formed *in vivo* under native conditions (Puig *et al*., 2001). However, the potential problem is that the efficiency of protein identification may be reduced due to the competition between the tagged protein and the corresponding endogenous protein to participate in a heterocomplex (Rohila *et al*., 2004; Abe *et al*., 2008). This problem might be solved by reducing endogenous proteins through RNAi or knockouts, or by overexpressing the tagged protein (Rohila *et al*., 2004). The other advantage of TAP-tagging is that in comparison to the single step purification approach, its two step purification can dramatically reduce the number of false positives and false negatives (Azarkan *et al*., 2007; O'Connor and Hames, 2008). On the other hand, the two successive steps of purification may fail to detect transient or weak interactions that may lost during the course of purifications.

1.2.2 Co-immunoprecipitation (Co-IP)

Co-IP is a classical method for protein-protein interaction analysis that has been used in numerous experiments (Phizicky and Fields, 1995). It is similar

to pull-down except that an antibody is used to purify the protein of interest instead of a tag (Lopper and Keck, 2007). In a Co-IP experiment, the antibody specific the antigen (protein of interest) is incubated with lysates which contains the antigen. If the antigen is immunoprecipitated with the antibody, then any protein that is associated with the antigen may be co-precipitated (Golemis and Adams, 2005). The immunoprecipitates are subsequently captured on protein A or G coated beads and washed to remove non-specifically bound proteins (Einarson *et al*., 2007). Finally, the components of the immunoprecipitates are eluted from the beads and analysed by Western blot, MS or sequencing.

Co-IP has been used be as a primary approach to identify novel interacting partners of a protein of interest. For instance, in order to identify proteins that bind inhibitor of apoptosis protein (IAP) homolog A (MIHA), Verhagen *et al*. (2000) transiently transfected 293T cells with FLAG-tagged mouse MIHA and coimmunoprecipitated associated proteins, and identified DIABLO (direct IAP binding protein with low pI) as an interactor of MIHA. But more often, Co-IP is used to corroborate the findings obtained from other large-sale high-throughput methods, such as the yeast two hybrid. The major advantage of Co-IP is that the proteins are present in their natural state of posttranslational modification, and both the antigen and the interacting proteins are present in the same relative concentrations as found in the cell (Phizicky and Fields, 1995). However, just as the affinity chromatography based methods, Co-IP may miss weak or transient interactions that can not survive the immunoprecipitation conditions (Vitale, 2002), or may identify false interactions because the production of the cellular lysate inevitably brings together the proteins that normally reside in different cellular compartments (Coates and Hall, 2002). It is also possible that the contact between the two proteins may mask the antigenic sites, or that an antibody which lacks sufficient specificity may recognize an irrelevant antigen. There are several criteria to be used to substantiate the authenticity of a Co-IP experiment. For example, confirm that the co-precipitated protein is obtained only by the

antibody against the target; verify that the antibody against the target antigen does not itself recognize the co-precipitated protein(s); determine that the interaction takes place in the cell and not as a consequence of cell lysis, etc. (Pierce, 2005).

1.2.3 Yeast two-hybrid systems

The first yeast two-hybrid system was described by Fields and Song (1989). So far, the yeast two-hybrid system is the most commonly used method for large-scale, high-throughput identification of novel protein-protein interactions (Walhout *et al*., 2000). As of 2004, the number of papers in which yeast two-hybrid system was used is about 10,000 (Golemis and Adams, 2005). In the past few years, a number of new two-hybrid systems, such as the Ras Recruitment Systems, Split-ubiquitin systems, and reverse two-hybrid systems, have been developed to improve the technology and suit different users (Causier, 2004). The conventional GAL4 yeast two-hybrid system (GAL4 system) and the Ras Recruitment Systems are described respectively in this section.

1.2.4.1 The GAL4 yeast two-hybrid system

The “Conventional” GAL4 yeast two-hybrid system exploits the fact that the intact Gal4 protein is a transcriptional activator that has two separate domains: a DNA binding domain (BD) and an activating domain (AD), and that the split protein does not work unless the two domains are in physical contact. Two plasmids, the bait containing plasmid and the prey containing plasmid need to be constructed. The bait, which is the protein of interest, is fused to the BD. The prey, such as a cDNA library, is fused to the AD. Plasmids are introduced into an appropriate yeast strain by co-transformation, sequential transformation, or by yeast mating (Causier, 2004). The interaction between the bait and prey will lead to physical contact of the BD and the AD, which subsequently induce reporter gene expression (Figure 1.8).

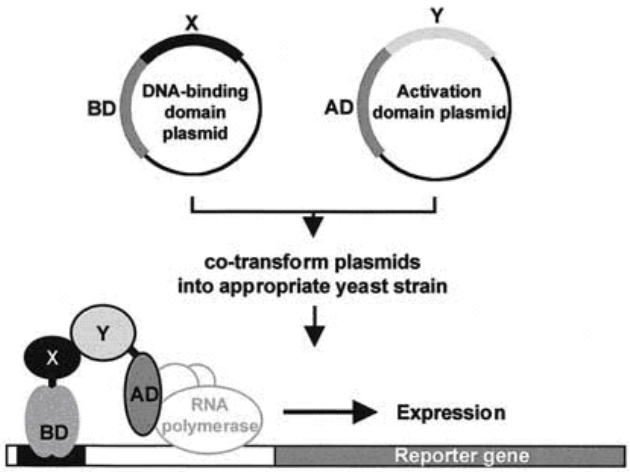


Figure 1.8. The principle of the GAL4 yeast two-hybrid system (taken from Causier and Davies, 2002). Interaction between a bait protein (X, fused in frame with BD) and a prey protein (Y, fused in frame with AD) reconstitutes an intact transcription factor that specifically binds to upstream elements of a reporter gene and activates its transcription.

Using the GAL4 system, numerous protein-protein interactions have been identified which have contributed greatly to current knowledge of the signal transduction networks in various organisms. However, the GAL4 system requires the transcriptional regulation that is induced in the nucleus of yeast by protein–protein interaction. It is known that membrane proteins tend to be insoluble and form aggregates when removed from the membrane (Milligan and White, 2001; Stagljar and Fields, 2002). Moreover, post-translational modifications (e.g. glycosylation and palmitoylation) of membrane proteins are unable to take place in the nucleus (Stagljar and Fields, 2002). Therefore, the conventional nuclear-based GAL4 system it not suited to analysis of membrane proteins. For example, interaction between GCR1 and GPA1 can not be detected by this system (Pandey and Assmann, 2004). For this reason, a few novel cytoplasm-based yeast two-hybrid systems, including the Ras Recruitment System, have been developed, which can overcome this problem while retaining the advantages of the conventional GAL4 system.

1.2.4.2 The Ras Recruitment System (RRS)

Ras proteins are small GTPases that regulate several signalling pathways controlling cell proliferation, differentiation and apoptosis (Lacal, 1997; Rebollo and Martinez-A, 1999; Crespo and Leon, 2000; Takai *et al*., 2001). The RRS developed by Aronheim's lab (Broder *et al*., 1998) is based on the strict requirement that Ras be localised to the plasma membrane for it to function. The yeast strain *cdc25-2* used in this system has a temperature sensitive mutation in Ras (yRas), so it grows only at permissive temperature 24°C, not at the restrictive temperature 36°C. The bait is fused to mammalian Ras (mRas) that lacks the CAAX box (it is required for Ras membrane targeting), therefore it is located in the cytoplasm. The prey, such as a cDNA library, is fused in frame with a myristoylation sequence that will lead to its membrane localisation. When the bait does not interact with the prey, mRas remains in the cytoplasm, thus yeast does not grow at 36°C. The interaction between the bait and prey would translocate the mRas to the membrane, which complements the mutation in yRas and enables yeast cells to grow at 36°C (Figure 1.9).

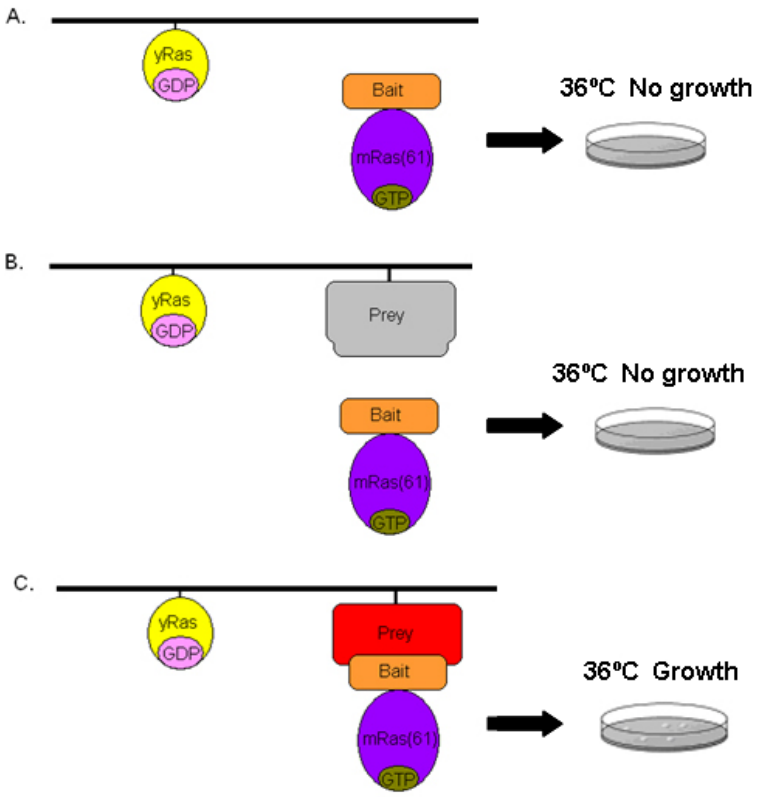


Figure 1.9 Schematic representation of the mechanism of the RRS (by project student Richard Browning, 2007; adapted from Broder *et al*., 1998). The bait is fused to mammalian Ras (mRas(61)) that lacks the CAAX box, thus it is not located on the plasma membrane. The prey is located on the plasma membrane through myristoylation (A). When there is no bait-prey interaction, mRas is localised in the cytoplasm, therefore yeast does not grow at the restrictive temperature 36°C (B). The interaction between the bait and prey would tanslocate mRas to the membrane so that the yeast grows at 36°C (C).

The RRS is a powerful system to detect protein-protein interactions at the inner face of the plasma membrane (Broder *et al*., 1998). It has been used to identify interactors for membrane proteins, and defining the protein domains involved in the interaction (Kohler and Muller, 2003). For instance, two MAGE proteins, necdin and MAGE-H1, were found to interact with the p75 neutrotrophin receptor in a RRS screen, and the domain responsible for the interaction was mapped to the intracellular “death” domain but not the “linker” and “tail” domains of p75 (Tcherpakov *et al*., 2002). Similarly, Polek *et al*. (2006) identified Ste20-related proline-alanine-rich kinase (SPAK) as a novel interactor of receptor expressed in lymphoid tissues (RELT, a type I transmembrane glycoprotein) using the RRS, and subsequently confirmed

the interaction by GST pull-down and Co-IP. It is worth noting that a reverse version of the RRS (rRRS), in which the bait is localised to the membrane and the prey is fused with the mRas, has been developed (Hubsman *et al*., 2001) and has been successfully used to identify protein-protein interactions with integral membrane proteins (Hubsman *et al*., 2001; Frankel *et al*, 2005). In addition, the RRS has also been used to study transcription factors that might show autoactivation in the GAL4 system. For example, using the RRS, Hennemann *et al*. (2003) identified Krim-1A and Krim-1B as interactors of c-Myc which is a transcription factor that belongs to nuclear phosphoproteins family, and successfully verified their interactions using Co-IP and a mammalian two-hybrid assay. In a work carried out by Takemaru *et al* (2003), a screening using the RRS with the Arm repeat 8 to the C-terminus of the β-catenin protein (contains 12 Arm repeats) as bait identified 79 positive clones, 59 of which encoded Chibby. The authors then used the RRS to define the Chibby binding domain of β-catenin and found that Arm repeat 10 to the C-terminus of β-catenin is required for its binding to Chibby. The result obtained from the RRS was supported by both pull-down and Co-IP (Takemaru *et al*., 2003). Altogether, the RRS is a valuable alternative to the conventional GAL4 system and can widely be used among protein classes for protein-protein interaction analysis (Kruse *et al*., 2006).

In general, the yeast two-hybrid system offers a number of advantages over other biochemical and genetic approaches for the analysis of protein-protein interactions. For example, the system is highly sensitive and relatively inexpensive (Causier and Davies, 2002). It detects *in vivo* protein-protein interactions without directly handling any protein molecules (Bockaert *et al*., 2004). Apart from the ability to screen libraries to isolate proteins that interact with a bait protein, it also allows for the analysis of known or suspected interactions. However, there are certain considerations that need to be taken into account when conducting the experiment. For example, a protein must be able to fold properly and exist as an active and stable fusion protein in yeast cells (Phizicky and Fields, 1995; Criekinge and Beyaert, 1999).

Interactions dependent on a posttranslational modification that does not occur in yeast cells will not be detected (Fields, 2001). In addition, preys may overcome the screening system by a mechanism independent of protein-protein interaction with the bait, so that false positives may be obtained when there is in fact no interaction present. For instance, a plant cell cycle gene might activate the RRS downstream of Ras. Furthermore, it is also possible that a third protein is bridging the two interacting protein partners (Criekinge and Beyaert, 1999). As a consequence, it is essential to verify putative protein-protein interactions using other methods.

Apart from the approaches addressed above, there are many other techniques that can be used in protein-protein interaction analysis. Each technique has strengths, but works under certain limitations. Choosing the most suitable one can greatly improve efficiency, and undoubtedly, a combination of different methods provides more reliable results. It is likely that the limitations of all these methods will drive the further invention and use of novel approaches in the future.

1.3 Research objectives

Although evidence to date indicates that GCR1 is the most promising candidate GPCR in the plant kingdom, the function of this protein is largely unknown. In order to achieve a better understanding of the GCR1 signalling pathway, this project aimed to use the RRS to screen an *Arabidopsis* cDNA library for proteins that interact with GCR1. Putative interactions generated from the library screening would be subsequently verified by rRRS and pull-down. The functional relevance of the putative interactors to GCR1 would be analysed using bioinformatics and genetic approaches. Figure 1.10 presents the flow chart that describes the main experimental procedures.

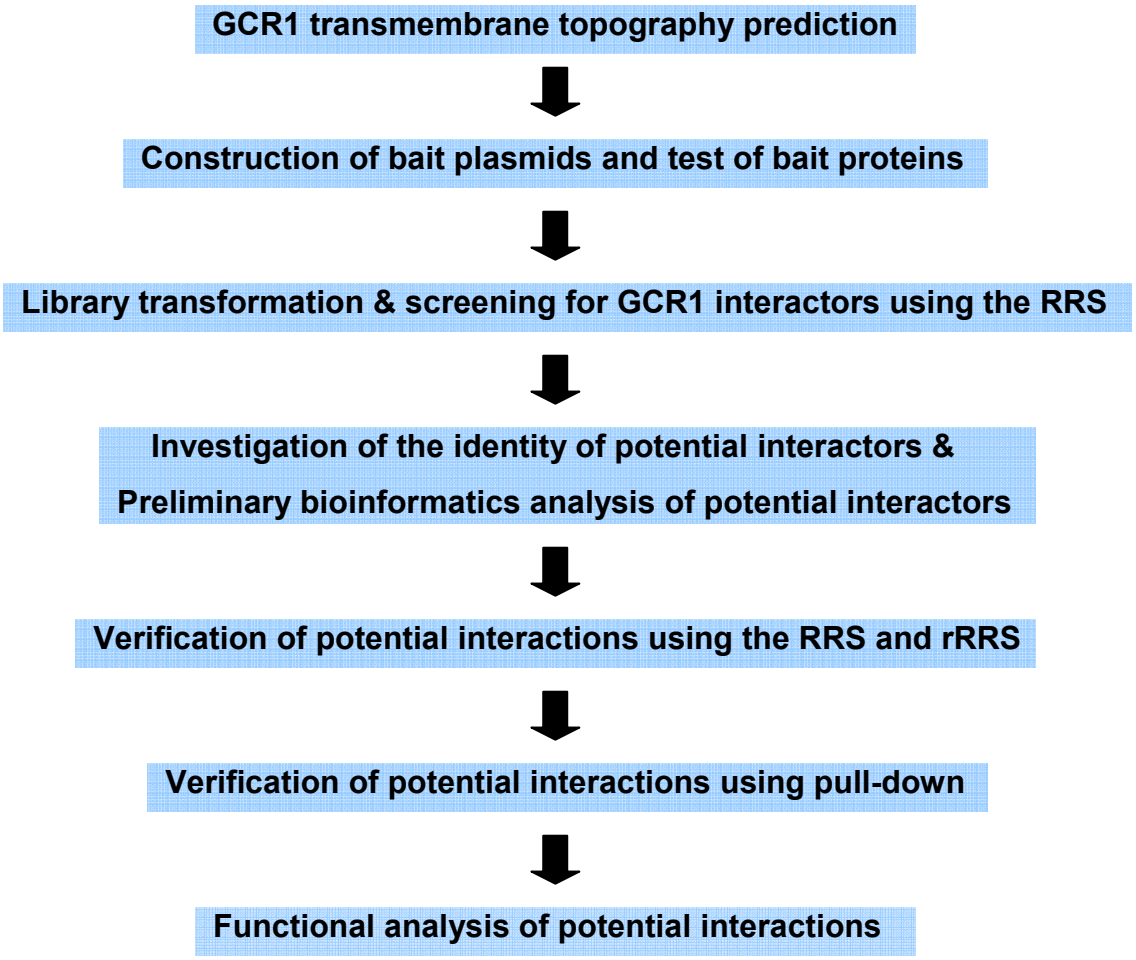


Figure 1.10 Flow chart of main experimental procedures

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Plasmids

2.1.1.1 Yeast expression plasmids

Yeast expression plasmids (Table 2.1) were a generous gift from Dr. Ami Aronheim (Technion-Israel Institute of Technology, Haifa, Israel) and are derivatives of those used in the RRS described by Broder *et al* (1998).

Table 2.1 Yeast expression plasmids

Plasmid	Abbr.	Description	Characteristics
pMet 425-Myc-Ras ∆ BamHI	pMetRas	RRS bait vector	Met425 promoter, amp^r, Myc tagged, encodes mRas & *Leu2*,
pUra-M ∆ polyA	pUra	RRS prey vector	Gal1 promoter, amp^r, encodes *Ura3*
pMet 425 ∆ BamHI	pMet	rRRS bait vector	Met425 promoter, amp^r, encodes *Leu2*,
pUra-Ras-RI	pUraRas	rRRS prey vector	Gal1 promoter, amp^r encodes mRas & *Ura3*,

pMet 425-Myc-Ras ∆ BamHI (referred henceforth as pMetRas, figure 2.1A and appendix 1) was used as the bait vector in the RRS. It contains the Met425 promoter, which allows expression of the bait only in the absence of methionine. Therefore the expression of bait is induced when yeast cells are grown on a medium lacking methionine and it is suppressed in the presence of methionine permitting elimination of false positives. It also contains a Myc epitope tag to facilitate detection of the bait fusion by Western blot using the anti-Myc antibody. Mammalian Ras (mRas) is fused to the Myc-tagged bait to complement the function of the mutated yeast Ras (yRas) in the RRS when

there is an interaction between the bait and prey. The vector also includes the *Leu2* gene to allow for selection of yeast cells containing the plasmid.

pUra-M Δ polyA (referred henceforth as pUra, figure 2.1B and appendix 1) was used as the prey vector in the RRS. It contains the Gal1 promoter, therefore prey expression is induced when yeast cells are grown on medium containing galactose and repressed in the presence of glucose. In addition, the vector contains a myristoylation sequence for prey membrane localization. The vector encodes the *Ura3* gene to allow for selection of yeast cells containing the plasmid.

pMet 425 Δ BamHI (referred henceforth as pMet, figure 2.1C and appendix 1) was used as the bait vector in rRRS. It contains the Met425 promoter, which allows expression of the bait only in the absence of methionine. Therefore the expression of bait is induced when yeast cells are grown on a medium lacking methionine and it is suppressed in the presence of methionine permitting elimination of false positives. The vector encodes the *Leu2* gene to allow for selection of yeast cells containing the plasmid.

pUra-Ras-RI (referred henceforth as pUraRas, figure 2.1D and appendix 1) was used as the prey vector in the rRRS. It contains the Gal1 promoter, therefore prey expression is induced when yeast cells are grown on medium containing galactose and repressed in the presence of glucose. The mRas is fused to the prey to complement the function of the mutated yRas in the rRRS when there is an interaction between the bait and prey. The vector encodes the *Ura3* gene to allow for selection of yeast cells containing the plasmid.

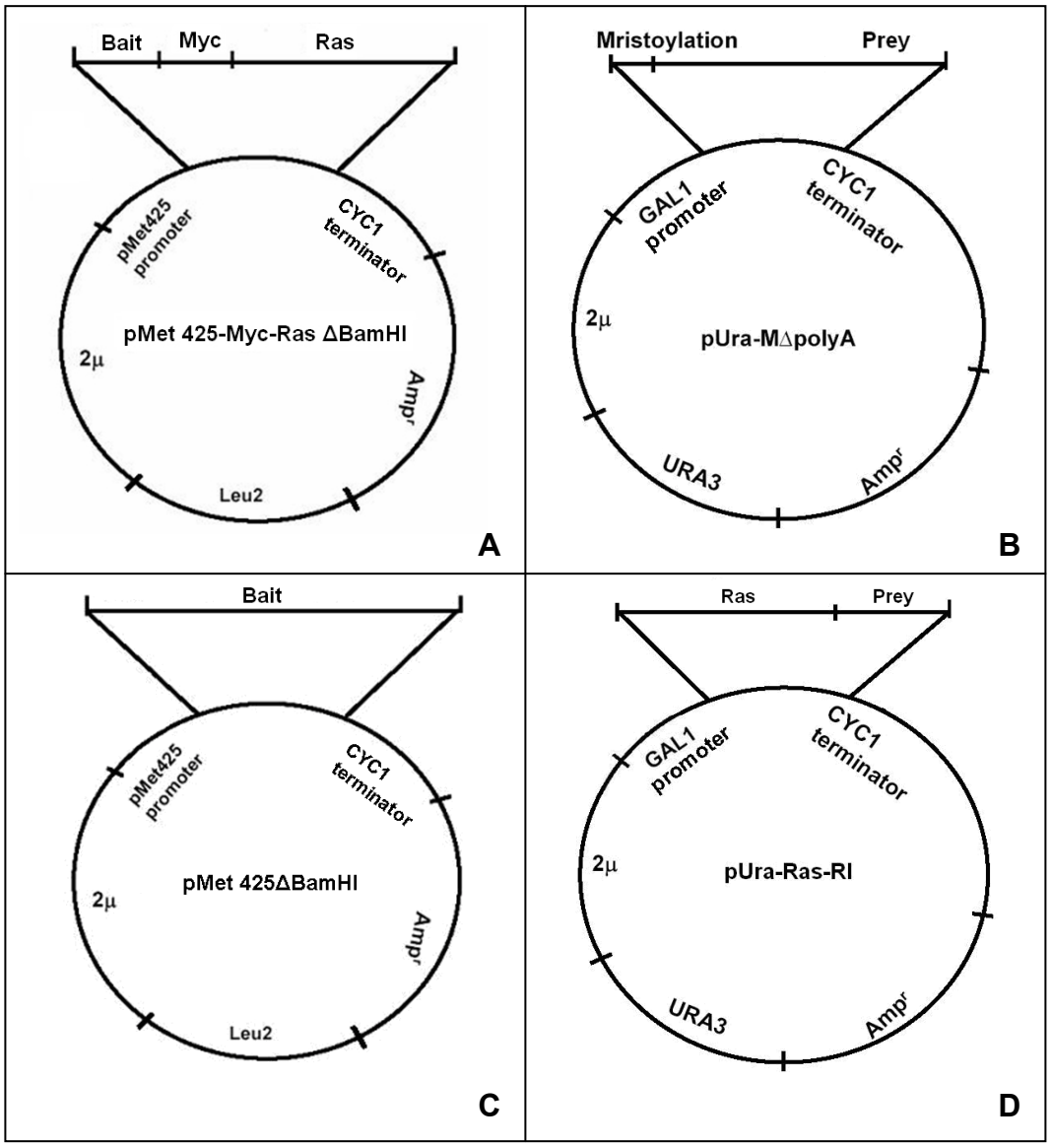


Figure 2.1 Plasmids used in yeast. In the RRS, pMet425-Myc-Ras Δ BamHI (A) was used as the bait vector and pUra-M Δ polyA (B) as the prey vector. In the rRRS, pMet425 Δ BamHI (C) was used as the bait vector and pUra-Ras-RI (D) was used as the prey vector.

2.1.1.2 *E.coli* expression plasmids

Three different *E.coli* expression plasmids were used in this project (Table 2.2).

Table 2.2 *E.coli* expression plasmids

Plasmid	Characteristics
pET22b(+)	T7 promoter, amp^r, an N-terminal *pelB* signal sequence, an optional C-terminal His tag
pET22b(+)myc	T7 promoter, amp^r, an N-terminal *pelB* signal sequence, an N-terminal Myc tag, an optional C-terminal His tag
pET28a(+)	T7 promoter, Kan^r, an N-terminal His tag, a thrombin cleavage site, an optional C-terminal His tag

The pET22b(+) vector (Novagen, figure 2.2 and appendix 1) has an N-terminal *pelB* signal sequence for potential periplasmic localization, plus an optional C-terminal His tag sequence.

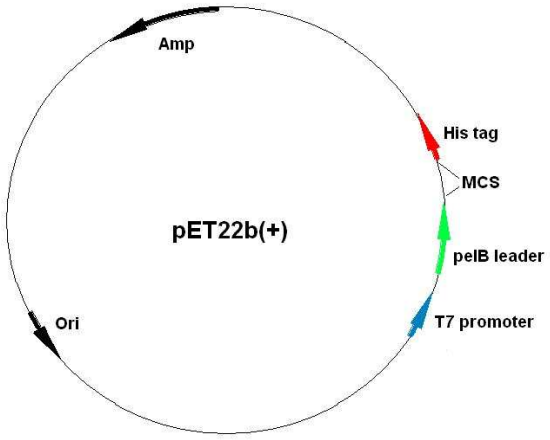


Figure 2.2 pET22b(+)

The pET22b(+)myc vector (Figure 2.3) was built by cloning the Myc sequence in frame to the *NcoI* site of pET22b(+). It has an N-terminal *pelB* signal sequence for potential periplasmic localization, followed by an N-terminal Myc tag sequence, plus optional C-terminal His tag sequence.

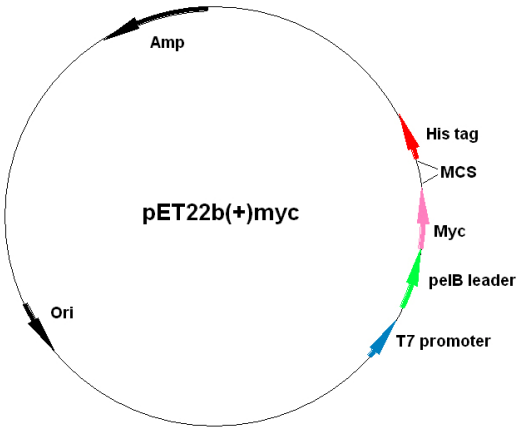


Figure 2.3 pET22b(+)myc

The pET28a(+) vector (Novagen, figure 2.4 and appendix 1) has an N-terminal His tag, a thrombin cleavage site to facilitate the removal of the His tag from the recombinant protein, plus an optional C-terminal His tag sequence.

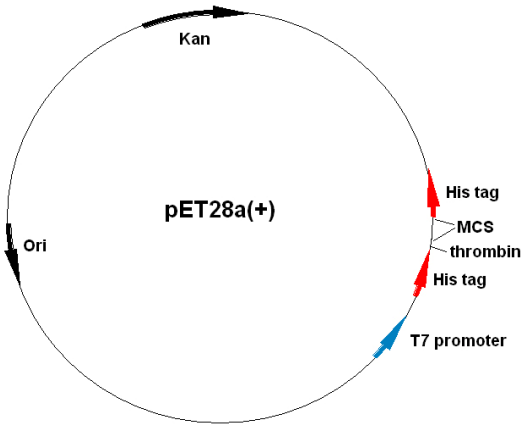


Figure 2.4 pET28a(+)

2.1.1.3 Cloning vectors

TOPO TA Cloning was used for the direct insertion of A-tailed PCR products into the pCR2.1-TOPO vector (Figure 2.5). It encodes the *lacZα* for blue-white color screening of colonies for ones with plasmids carrying inserts.

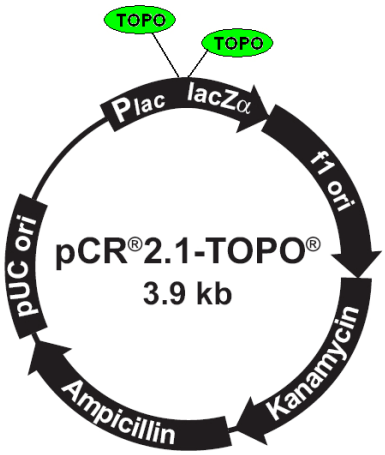


Figure 2.5 pCR2.1-TOPO cloning vector (Adapted from the user manual for TOPO TA Cloning®, Invitrogen).

The pENTR/D-TOPO Gateway entry vector (Figure 2.6) facilitates rapid, directional TOPO cloning of blunt-end PCR products for entry into the Gateway System (Earley *et al*., 2006)

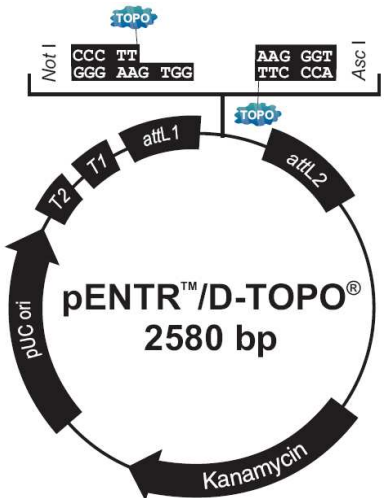


Figure 2.6 pENTR/D Gateway entry vector (adapted from the user manual for pENTR™ Directional TOPO® Cloning Kits, Invitrogen).

2.1.1.4 Plant expression plasmids

The pEarleyGate binary vectors (Figure 2.7) were derived from pFGC5941, which was built using a pCAMBIA plasmid backbone (Earley *et al*., 2006). The Gateway cassettes in each vector include attR1, a chloramphenicol

resistance gene (CmR), the ccdB killer gene and attR2. 35S, the cauliflower mosaic virus 35S promoter and its upstream enhancer. OCS, the 3' sequences of the octopine synthase gene, including polyadenylation and presumptive transcription termination sequences. BAR, the Basta herbicide resistance gene for selection of transgenic plants. Km, the bacterial kanamycin resistance gene within the plasmid backbone for plasmid selection in *E. coli* and *Agrobacterium tumefaciens*. Different pEarleyGate vectors allow engineering and expression of proteins fused in frame with HA, FLAG, cMyc or AcV5 tags at the amino-terminal end of the target proteins (Earley *et al*., 2006).

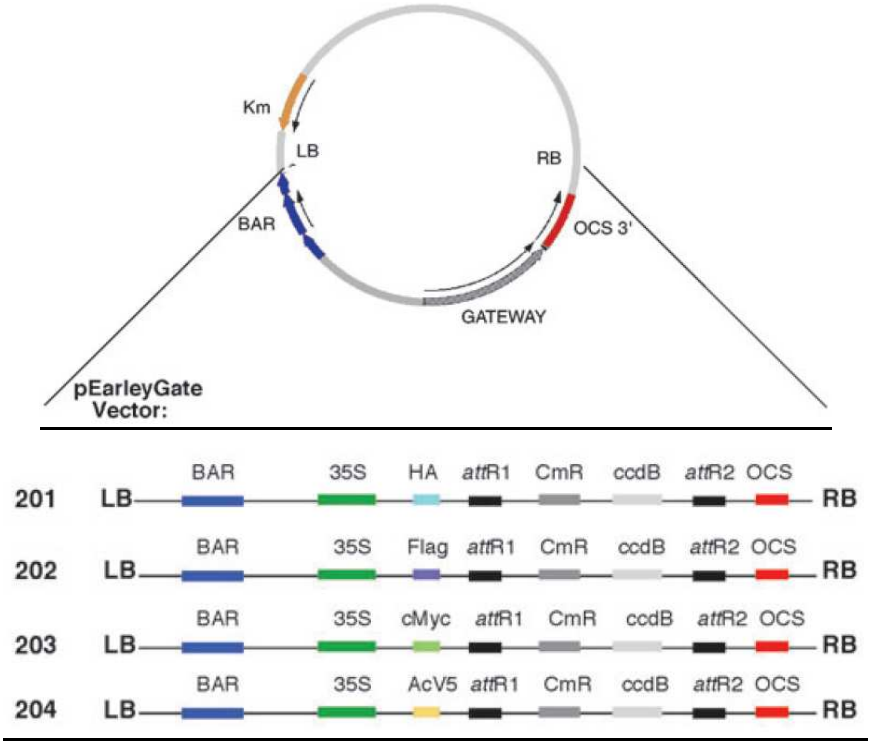


Figure 2.7 pEarleyGate vectors (adapted from Earley *et al*., 2006)

2.1.2 Yeast, *E.coli* and *Agrobacterium tumefaciens* strains

2.1.2.1 Yeast strain

The temperature-sensitive yeast strain *cdc25-2* (CDC25-2α ura3, lys2, leu2, trp1, hisΔ200, ade2-101) used in this project was a generous gift from Dr. Ami Aronheim (Technion-Israel Institute of Technology, Haifa, Israel). The yRas is rendered inactive at the restrictive temperature 36°C due to it lacking a functional Cdc25 guanyl nucleotide exchange factor. The yeast cells grow at 24°C, but not at the restrictive temperature 36°C.

2.1.2.2 *E.coli* strains

The TAM1 strain (Activemotif) was used for cloning purposes. Its genotype is: *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 (*ara-leu*)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G.

The *E.coli* strains (Novagen) used for protein expression purposes are listed in table 2.3.

Table 2.3 *E.coli* strains used for protein expression purposes

Strain	Genotype	Description/Application
BL21(DE3)	F^- *ompT* $hsdS_B(r_B^-m_B^-)$*gal dcm* (DE3)	General purpose expression host.
Rosetta2 (DE3)	F^- *ompT* $hsdS_B(r_B^-m_B^-)$*gal dcm* (DE3) pRARE2[7](Cam^R)	Provides seven rare codon tRNAs
Origami (DE3)	*Δara–leu7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL* F'*[lac+ lacIq pro]* (DE3) *gor522 ::*Tn*10 trxB* (KanR, StrR, TetR)	Two mutations in cytoplasmic disulfide reduction pathway enhance disulfide bond formation in *E. coli* cytoplasm.
Rosetta-gami (DE3)	*Δara–leu7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL* (DE3) F'*[lac+ lacIq pro] gor522 ::*Tn*10 trxB* pRARE6 (CamR, KanR, StrR, TetR)	Two mutations in cytoplasmic disulfide reduction pathway enhance disulfide bond formation in *E. coli* cytoplasm. Provides six rare codon tRNAs.

2.1.2.3 *Agrobacterium tumefaciens* strain

Agrobacterium tumefaciens strain GV3101 (Rifampicin, Gentamycin and Ampicillin resistant) was used in plant transformation experiments.

2.1.3 cDNA library for library screening

A cDNA library which contains 8.5 x 10^5 primary clones was prepared from mRNA isolated from *Arabidopsis* 10 days old seedlings and organs from mature flowering plants including leaves, stems, flowers, roots and siliques. Double stranded cDNAs were synthesised using Stratagene's cDNA synthesis kit and directionally cloned into the RRS prey vector pUra by Dr. Baoxiu Qi. The library was subsequently re-amplified (section 2.2.6) and used in the library screening.

2.1.4 Yeast, *E.coli* and *Agrobacterium tumefaciens* growth media

Media used in the project are listed in table 2.4. Detailed composition of media is in appendix 2.

Table 2.4 Yeast, *E.coli* and *Agrobacterium tumefaciens* growth media

Application	Full medium name	Abbr.
For yeast	YPD	YPD
	2xYPD	2xYPD
	Glucose-leucine-methionine	Glu-L-M
	Glucose-leucine+methionine	Glu-L+M
	Glucose-leucine+4methionine	Glu-L+4M
	Glucose-uracyl+methionine	Glu-U+M
	Galactose-uracyl+methionine	Gal-U+M
	Glucose-leucine-uracyl-methionine	Glu-L-U-M
	Glucose-leucine-uracyl+methionine	Glu-L-U+M
	Galactose-leucine-uracyl-methionine	Gal-L-U-M
	Galactose-leucine-uracyl+4methionine	Gal-L-U+4M

For *E.coli*	LB	LB
	NZY^+	NZY^+
	SOC	SOC
For *Agrobacterium*	2xYT	2xYT

2.1.5 Plant material

Arabidopsis thaliana: Wild type Columbia-0 (Col-0) and Salk insertion lines were obtained from the Nottingham Arabidopsis Stock Center (NASC, UK)
Tobacco: *Nicotiana benthamiana*.

2.1.6 Reagents and kits

Reagents and kits used are described in methods.

2.2 Methods

2.2.1 Protein transmembrane topography prediction programs

The following online protein membrane topography prediction programs were used to facilitate GCR1 and *zf*-DHHC1 membrane topography prediction.

2.2.1.1 TMpred – Transmembrane Prediction

Three different settings were used, with 17 amino acids (aa), 19 aa and 21 aa as the minimum length of the hydrophobic part of the transmembrane helix respectively.
http://www.ch.embnet.org/software/TMPRED_form.html

2.2.1.2 TMHMM – Tied-mixture Hidden Markov Modeling

Default settings were used.
http://www.cbs.dtu.dk/services/TMHMM-2.0

2.2.1.3 TMAP (single)

Default settings were used.

http://bioinfo4.limbo.ifm.liu.se/tmap/single.html

2.2.1.4 TopPred

Default settings were used.

http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=tmap

2.2.2 Gene cloning

The general gene cloning methods are described in this section. The constructs made in this project are listed in table 2.5 – 2.8. The primer sequences are listed in appendix 3.

Table 2.5 Constructs made for use in the RRS

	Construct name	Primers for amplifying the insert
Bait Constructs	pMetRas-i1	LP1FXmaI + LP1RXmaI
	pMetRas-i2	LP2FXmaI + LP2RXmaI
	pMetRas-i1-i2	GCR1bgn+ LP1R LP2F + LP3RXamI LP1FXmaI + LP2RXmaI
	pMetRas-i3	LP3FXmaI + LP3RXmaI
	pMetRas-Cter	CterFXmaI + CterRXmaI
	pMetRas-VLCYCLF-i1-i2	VLP1FXmaI + LP2RXmaI
	pMetRas-CYCLF-i1-i2	CLP1FXmaI + LP2RXmaI
	pMetRas-VLCYCLF-i1	VLP1FXmaI + LP1RXmaI
	pMetRas-CYCLF-i1	CLP1FXmaI + LP1RXmaI
	pMetRas-i1-GGG-i2	GCR1bgn + LP1GGGrev GGGLP2for + LP3RXmaI LP1FXmaI + LP2RXmaI
Prey Constructs	pUra-TRX3Ser39 (pUra-S39)	TRX3BegBmRI + TRX3^{Ser39}R AtTRX3^{Ser39}F + TRX3EndRI TRX3BegBmRI + TRX3EndRI
	pUra-TRX3Ser42 (pUra-S42)	TRX3BegBmRI + TRX3^{Ser42}R TRX3^{Ser42}F + TRX3EndRI TRX3BegBmRI + TRX3EndRI
	pUra-TRX3Ser39Ser42 (pUra-S39S42)	TRX3BegBmRI + TRX3^{Ser42}R TRX3BegBmRI + TRX3EndRI
	pUra-TRX4Ser40 (pUra-S40)	TRX4BegBmRI + TRX4^{Ser40}R TRX4^{Ser40}F+ TRX4EndRI TRX4BegBmRI+ TRX4EndRI
	pUra-TRX4Ser43 (pUra-S43)	TRX4BegBmRI + TRX4^{Ser43}R TRX4^{Ser43}F+ TRX4EndRI TRX4BegBmRI + TRX4EndRI
	pUra-TRX4Ser40Ser43 (pUra-S40S430)	TRX4BegBmRI + TRX4^{Ser43}R TRX4BegBmRI + TRX4EndRI

Table 2.6 Constructs made for use in the rRRS

	Construct name	Primers for amplifying the insert
Bait Constructs	pMet-Nter-i2	GCR1BegHindIII + LP2EndHindIII
	pMet-i3-Cter	LP3BegHindIII + CterEndHindIII
	pMet-Nter-Cter	GCR1BegHindIII + CterEndHindIII
	pMet-*zf*-DHHC1*	ZFBegHindIII + ZFEnd
Prey Constructs	pUraRas-TRX3	TRX3BegBmRI + TRX3EndRI
	pUraRas-TRX4	TRX4BegBmRI + TRX4EndRI
	pUraRas-VLCYCLF-i1	VLP1FEcoRI + LP1EndEcoRI
	pUraRas-CYCLF-i1	CLP1FEcoRI + LP1EndEcoRI
	pUraRas-i1	LP1FEcoRI + LP1EndEcoRI
	pUraRas-i2	LP2FEcoRI + LP2EndEcoRI
	pUraRas-i1-i2	LP1FEcoRI + LP2EndEcoRI
	pUraRas-i1-GGG-i2	LP1FEcoRI + LP2EndEcoRI

**zf*-DHHC1 was initially cloned into pCR2.1 using the TOPO cloning method (section 2.2.1.6), and then cloned into pMet using normal cloning method.

Table 2.7 Constructs made for use in the pull-down assay

	Construct name	Primers for amplifying the insert
Bait Constructs	pET28a(+)-CYCLF-i1	CYCLP1BegEcoRI + LP1EndEcoRI
	pET28a(+)-i1	LP1BegEcoRI + LP1EndEcoRI
	pET28a(+)-i2	LP2BegEcoRI + LP2EndEcoRI
	pET28a(+)-i1-GGG-i2	LP1BegEcoRI + LP2EndEcoRI
	pET28a(+)-i3	LP3BegHindIII 2 + LP3EndHindIII
	pET28a(+)-Cter	CterBegHindIII 2 + CterEndHindIII
	pET22b(+)-Nter-i2	GCR1begHindIII 2 + LP2EndHindIII
	pET22b(+)-i1-GGG-i2*	N/A
	pET22b(+)-i3*	N/A
Prey Constructs	pET22b(+)myc-TRX3	TRX3BegEcoRI+ TRX3EndRI
	pET22b(+)myc-TRX4	TRX4BegEcoRI+ TRX4EndRI
	pET22b(+)myc-TRX3S42	TRX3BegEcoRI+ TRX3EndRI
	pET22b(+)myc-TRX4S43	TRX4BegEcoRI+ TRX4EndRI
	pET22b(+)myc-*zf*-DHHC1 *	N/A

* The inserts of these constructs were prepared by restriction digestion of the pMetRas-i1-GGG-i2, pMetRas-i3 and pCR2.1-*zf*-DHHC1 constructs respectively.

Table 2.8 Constructs made for use in the Co-IP

Construct name	Primers for amplifying the insert
pENTR/D-GCR1*	GCR1CACCbeg + CterEndHindIII
pENTR/D-GCR1-FLAG*	GCR1CACCbeg + FLAGIC
pENTR/D-TRX3*	TRX3CACCbeg + TRX3EndRI
pENTR/D-TRX4*	TRX4CACCbeg + TRX4EndRI
pENTR/D-TRX3Ser42 (pENTR/D-S42)*	TRX3CACCbeg + TRX3EndRI
pENTR/D-TRX4Ser43 (pENTR/D-S43)*	TRX4CACCbeg + TRX4EndRI
pENTR/D-*zf*-DHHC1*	ZFCACCbeg + ZFEnd
pEarleyGate201-GCR1 (HA-GCR1)**	N/A
pEarleyGate202-GCR1 (FLAG-GCR1)**	N/A
pEarleyGate101-GCR1 or pCambia1300-GCR1 (GCR1-FLAG)**	N/A
pEarleyGate203-GCR1 (Myc-GCR1)**	N/A
pEarleyGate204-GCR1 (AcV5-GCR1)**	N/A
pEarleyGate203-TRX3 (Myc-TRX3)**	N/A
pEarleyGate203-TRX4 (Myc-TRX4)**	N/A
pEarleyGate203-TRX3Ser42 (Myc-S42)**	N/A
pEarleyGate203-TRX4Ser43 (Myc-S43)**	N/A
pEarleyGate202-TRX4 (FLAG-TRX4)**	N/A
pEarleyGate203-*zf*-DHHC1 (Myc-*zf*-DHHC1)**	N/A
pEarleyGate202-*zf*-DHHC1 (FLAG-*zf*-DHHC1)**	N/A

* These constructs were made using the TOPO cloning method (section 2.2.2.6)
**These constructs were made using the LR recombination method (section 2.2.2.7)

2.2.2.1 Preparation of insert

The inserts were amplified by PCR using either the KOD DNA polymerase (Novagen) to generate blunt ended products (Table 2.9a and 2.9b) or the Expand High Fidelity^{Plus} PCR System (Roche) to generate A-tailed products (Table 2.10a and 2.10b).

Table 2.9a KOD PCR – reaction mix

Volume	Reagents
5 µl	Template DNA
5 µl	10x Buffer#1 (for KOD DNA Polymerase)
5 µl	dNTPs (final concentration 0.2 mM)
2 µl	$MgCl_2$ (final concentration 1 mM)
5 µl	Forward primer (3 pmol/µl)
5 µl	Reverse primer (3 pmol/µl)
0.5 µl	KOD DNA Polymerase (2.5 U/µl)
22.5 µl	ddH_2O

Table 2.9b KOD PCR – programme

	Steps	Temperature (°C)	Time
1	Denaturation	98	15"
2	Annealing	55	5"
3	Extension	72	20"
4	Repeat (1-3)	-	24 cycles
5	Final polishing	72	6'

Table 2.10a Expand High Fidelity^{Plus} PCR – reaction mix

Volume	Reagents
5 µl	Template DNA
10 µl	Expand High Fidelity^{PLUS} Reaction Buffer (5x) with 7.5 mM $MgCl_2$
0.5 µl	dNTPs (final concentration 0.2 mM)
5 µl	Forward primer (3 pmol/µl)
5 µl	Reverse primer (3 pmol/µl)
0.5 µl	Expand High Fidelity^{PLUS} Enzyme Blend (5 U/µl)
24 µl	ddH_2O

Table 2.10b Expand High Fidelity^{Plus} PCR – programme

	Steps	Temperature (°C)	Time
1	Denaturation	94	3'
2	Denaturation	94	15''
3	Annealing	55	30''
4	Extension	72	1'
5	Repeat (2-4)	-	9 cycles
6	Denaturation	94	15''
7	Annealing	55	30''
8	Extension	72	1'3''
9	+3 sec/cycle	-	20''
10	Repeat (6-9)	-	19 cycles
11	Final polishing	72	6'

PCR products were analyzed by electrophoresis through agarose gels. 2% gel (Agarose MS, Roche) was used for inserts less than 150 bp. 1% gel (Agarose, Invitrogen) was used for inserts more than 150 bp. The size of the PCR products was determined by comparison with the 50 bp and 100 bp DNA ladder (Invitrogen), Hyperladder II (Bioline) or Ultra low ladder (Fermentas).

After purification (Roche High Pure PCR Product Purification Kit (Roche) or QIAEX II Gel Extraction Kit (Qiagen)), PCR products were digested with respective restriction enzyme at 37°C for 2 – 4 hours (Table 2.11). Digested DNA was extracted with an equal volume of phenol:chloroform (1:1 v/v), then with an equal volume of chloroform. The DNA was ethanol precipitated by adding 1/10 volume of 3 M NaAcetate (pH 5.2) and 2.5 volume of ice cold 100% ethanol. DNA was incubated at –20°C overnight or at –80°C for 2 hours to facilitate DNA precipitation. The sample was centrifuged at 16,000 x *g* in a 1.5 ml centrifuge tube for 1 hour at 4°C. The pellet was washed with 0.5 ml of ice cold 70% ethanol, centrifuged as above for 15 min, and air dried. The DNA was finally resuspended in 6 µl ddH_2O. The DNA concentration and quality was checked by gel electrophoresis as described above.

Table 2.11 A typical digestion reaction

Volume	Reagents
5 µl	10x Buffer
0.5 µl	BSA
2.5 µl	Restriction enzyme
40 µl	DNA
2 µl	ddH_2O

2.2.2.2 Preparation of vector

The vector was digested with respective restriction enzyme at 37°C for 3 hours or overnight as described in table 2.11. Shrimp Alkaline Phosphatase (Promega) was added to the reaction mixture and incubated at 37°C for 30 min to de-phosphorylate the vector ends to prevent the vector from self re-ligation. The digested DNA was extracted with 1:1 phenol/chloroform, precipitated by ethanol as described above, and resuspended in 20 µl ddH_2O. The quality of the DNA was checked by gel electrophoresis as described above. The concentration of the plasmid DNA was measured by spectrophotometry at A260nm using a spectrophotometer (Spectronic Uicam).

2.2.2.3 Ligation and transformation of competent *E.coli* cells

The digested and purified insert and vector were ligated using T4 DNA ligase (Promega). A typical ligation reaction (Table 2.12) was carried out at room temperature for 1-2 hours or at 16°C overnight.

Table 2.12 A typical ligation reaction

Volume	Reagents
2.5 µl	2X T4 DNA ligase buffer
0.5 µl	T4 DNA ligase
3 µl*	Insert
2 µl*	Vector

*The insert and vector were mixed and freeze-dried to 2 µl.

5 µl ligation mixture was mixed with 25 µl competent *E.coli* cells and incubated on ice for 30 min. After heat-shock at 42°C for 1 min, 125 µl SOC medium was added immediately and mixed gently with the cells. Cells were allowed to recover for 1 hour at 37°C before plated onto LB plate containing respective antibiotics (100 µg/ml ampicillin or 50 µg/ml kanamycin) and incubated at 37°C overnight.

2.2.2.4 Colony PCR

Colony PCR was carried out using a vector specific forward primer, and insert specific reverse primer to screen for *E.coli* colonies that contained the insert with the correct orientation. A sample of colony to be screened was transferred with a toothpick and resuspended in 15 µl TE buffer containing 0.6 µl Proteinase K (5 mg/ml, Roche). The mixture was incubated at 55°C for 15 min in order to release the plasmids from *E.coli* cells followed by 80°C for 15 min to denature the proteinase K. 1 µl of the digest was used as template in a 10 µl of PCR. The PCR mix and programme are listed in table 2.13a and 2.13b respectively.

Table 2.13a A typical colony PCR mix

Volume	Reagents
1 µl	DNA template
1 µl	5' primer (3µM)
1 µl	3' primer (3µM)
5 µl	2x PCR premix (ABgene)
2 µl	ddH_2O

Table 2.13b A typical colony PCR – programme

	Steps	Temperature (°C)	Time
1	Denaturation	94	3'
2	Denaturation	94	15"
3	Annealing	55	30"
4	Extension	72	1'
5	Repeat (2-4)	-	9 cycles
6	Denaturation	94	15"
7	Annealing	55	30"
8	Extension	72	1'3"
9	+3 sec/cycle	-	20"
10	Repeat (6-9)	-	19 cycles
11	Final polishing	72	6'

The colonies with insert of the correct size and orientation were patched on LB plates containing appropriate antibiotics and incubated at 37°C overnight for plasmid isolation as described below.

2.2.2.5 Preparation of plasmid and sequencing

Plasmids were extracted and purified using the Wizard *Plus* SV Minipreps DNA Purification System (Promega) according to manufacturer's instructions. 3 µl of the plasmid was digested with the relevant restriction enzyme in a 10 µl reaction to further confirm that the plasmid was the correct construct. The concentration of the plasmid DNA was measured by spectrophotometry at

A260nm. 1 µg of the plasmid DNA and 30 pmoles of sequencing primer were used for sequencing (Lark Technologies DNA sequencing services).

2.2.2.6 TOPO-cloning

The insert was amplified by PCR using either the KOD proof-reading DNA polymerase to generate blunt ended product or the Expand High Fidelity PCR System to generate A-tailed product as described in 2.2.1.1. A typical reaction (Table 2.14) was carried out at room temperature for 1 – 3 hours.

Table 2.14 A typical TOPO cloning reaction

Volume	Reagents
1 µl	1:10 dilution of PCR product
0.5 µl	Salt
0.5 µl	Vector (pCR2.1, or pENTR/D)

The reaction mixture was transformed with 25 µl competent *E.coli* cells as described in 2.2.1.3, plated out on respective LB plates (50 µg/ml kanamycin, 35 µg/ml X-gal and 12 µg/ml IPTG for pCR2.1; or 50 µg/ml kanamycin for pENTR/D) and incubated at 37°C overnight.

2.2.2.7 LR recombination

The entry clone in pENTR/D was digested with the restriction enzyme *HpaI* (NEB) at 37°C for 3 – 4 hours and heated at 65°C for 20 min to deactivate the enzyme. A typical LR recombination reaction was carried out at 25°C for 3 hours or overnight (Table 2.15).

Table 2.15 A typical LR recombination reaction

Volume	Reagents
2 µl (100 ng)	Entry clone (pENTR/D)
1 µl (100 ng)	Destination vector (pEarleyGate)
1 µl	LR Clonase II enzyme mix
1 µl	ddH_2O

2.2.3 Yeast transformation

One yeast colony was inoculated into 10 ml YPD liquid medium and grown at 24°C overnight with shaking. 1 ml of yeast cells were harvested by centrifugation at 1,000 x *g* for 1 min. The supernatant was discarded. Salmon sperm DNA (ssDNA) (10 mg/ml, Sigma D-1626) was boiled for 10 min and immediately cooled on ice for 5 – 10 min. 1 µg of plasmid was mixed together with 10 µl of preboiled ssDNA. Then 0.5 ml of LiPEG (40% PEG_{3350} in 100 mM LiAc containing TE) was added and vortexed. The mixture was kept on the bench overnight and plated onto appropriate plates. Cells were grown at 24°C for 5 – 7 days until colonies appeared.

2.2.4 Western blot

2.2.4.1 Sample preparation

(1) Sample preparation for yeast cells

Yeast cells were grown in 10 ml appropriate medium at 24°C with shaking for 24 hours. A600nm of the cell culture was measured using a spectrophotometer. 5 ml of yeast cell at an OD of 1.0 were harvested by centrifugation at 16,000 x *g* for 2 min. The supernatant was discarded and the pellet was resuspended in 100 µl ddH_2O. 100 µl of 0.2 M NaOH was added and incubated at room temperature for 5 min. Cells were pelleted again as above, and resuspended in 50 µl sample buffer (0.06 M Tris, 5% v/v Glycerol, 2% SDS, 0.0025% Bromophenol blue, 5% v/v β-mercaptoethanol). The

sample was boiled for 3 min, and centrifuged at 16,000 x *g* for 3 min before loading. 10 µl of the supernatant was used in the SDS – PAGE.

(2) Sample preparation for *E.coli* cells

Please refer to section 2.2.13.2 for details.

(3) Sample preparation for tobacco leaves

Tobacco leaf discs (~2cm^2) were ground in 100 µl of sample buffer in a 1.5 ml centrifuge tube using a blue plastic pestle, boiled for 5 min and centrifuged at 16,000 x *g* for 5 min before loading. 15 µl of the supernatant was used in the SDS – PAGE.

2.2.4.2 SDS-PAGE gel preparation

The SDS-PAGE gel (8 cm x 7.3 cm x 0.075 cm) was made of a resolving gel containing 12% of acrylamide and 0.1% SDS (pH 8.8) (Table 2.16a) and a stacking gel containing 4% acrylamide and 0.1% SDS (pH 6.8) (Table 2.16b) using the Mini-PROTEAN 3 Cell (Bio-Rad).

Table 2.16a Resolving gel solution

12% Resolving gel (10 ml)	
Chemical	Volume
40% Acrylamide	3 ml
2% Bis-Acrylamide	1.6 ml
ddH_2O	2.75 ml
1.5 M Tris-HCl pH8.8	2.5 ml
10% SDS	100 µl
10% APS	50 µl
TEMED	5 µl

Table 2.16b Stacking gel solution

4% stacking gel (5 ml)	
Chemical	Volume
40% Acrylamide	485 µl
2% Bis-Acrylamide	250 µl
ddH_2O	3.75 ml
1.0 M Tris-HCl pH6.8	625 µl
10% SDS	50 µl
10% APS	25 µl
TEMED	5 µl

2.2.4.3 Electrophoresis

The gel was run at 80V for 30 min and 100V for 70 min in TGS buffer (25 mM Tris, 192 mM glycine and 0.1% SDS, pH8.3) using the Mini-PROTEAN 3 Cell electrophoresis system (Bio-Rad). The gel was removed from the glass plates. Polypeptides were detected by staining the gel in Coomassie blue (0.1% Coomassie Brilliant Blue R-250, 10% acetic acid and 25% methanol) for 1 hour, and destaining (10% acetic acid and 25% methanol) overnight.

2.2.4.4 Western blot

A fresh gel was prepared and the proteins were separated as described above. The proteins were transferred from the gel onto the membrane (Immobilon – P Transfer Membrane, Millipore) by electrophoretic transfer in transfer buffer (25 mM Tris, 192 mM glycine and 15% methanol) for 1 hour at 100V using the Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad). The membrane was removed from the cell and incubated in 50 ml blocking buffer (0.3% Casein, 500 mM NaCl and 20 mM Tris, 0.1% Tween 20, pH7.4) with shaking for 1 hour. It was incubated, sealed in a plastic bag from which all air bubbles were squeezed out, with 3 ml 1/5,000 anti-Myc antibody (Invitrogen) (or 1/5,000 anti-His antibody (Novagen), 1/10,000 anti-ECS antibody (Bethyl), 1/10,000 anti-HA antibody (Bethyl), 1/10,000 anti-AcV5 antibody (Bethyl), 1/100,000 ~ 1/10,000 anti-cMyc antibody (Benthyl)) with shaking overnight. It was removed from the bag and washed for 10 min 3 times in blocking buffer before incubated in 30 ml 1/30,000 goat anti-mouse or goat anti-rabbit IgG (Alkaline phosphatase conjugated, Sigma) for 1.5 hours. The membrane was washed 3x 10 min in blocking buffer and 3x 10 min in AP buffer (100 mM Tris, 100 mM NaCl, 50 mM $MgCl_2$, pH9.5). Immuno-stained polypeptides were detected by incubating in 20 ml of premixed 0.56 mM BCIP (5-bromo-4-chloro-3-indolyl phosphate) and 0.48 mM NBT (nitro blue terrazolium) for 10-20 min. Membrane was washed with ddH_2O several times and dried. Photos were taken for record.

2.2.5 Bait autoactivation test

Bait containing yeast colonies were streaked onto Glu-L+M (for the RRS) or Glu-U+M (for the rRRS) plate and grown at 24°C for 3 days. The colonies were replica plated onto two Glu-L+M (for the RRS) or Glu-U+M (for the rRRS) plates. One was grown at 36°C for 3 days, and the other at 24°C as control. Colonies that grew at 24°C but not at 36°C were selected and checked in an autoactivation test. The autoactivation test was performed by replica plating the yeast cells from Glu-L+M (for the RRS) or Glu-U+M (for the rRRS) plate onto two Glu-L-M plates (for the RRS) or Gal-U+M (for the rRRS) plates, one grown at 36°C for 3-5 days, the other at 24°C as control.

2.2.6 Reamplification of an *Arabidopsis* cDNA library

100 µl of XL10-Gold ultracompetent cells (Stratagene) were thawed on ice and mixed with 4 µl of β-mercaptoethanol in a prechilled 15 ml Flacon 2059 polypropylene tube. The mixture was incubated on ice for 10 min with gent swirling every 2 min. 1 µg of the primary cDNA library was added to the cells and incubated on ice for 30 min. The tube was heat-pulsed in a 42°C water bath for 30 sec and then incubated on ice for 2 min. 0.9 ml of preheated (42°C) NZY^+ broth was added to the tube and incubated at 37°C for 1 hour with shaking at 250 rpm. The cells were plated out onto 30 individual LB-amp plates (14 cm) with about 35 µl on each plate and incubated at 37°C overnight. 5ml of liquid LB was added to each plate and a spreader was used to gently scrub the plate to collect the cells. The cells were finally transferred to several 50 ml tubes. Plasmids were extracted and purified using the Wizard *Plus* SV Midipreps DNA Purification System (Promega) according to manufacturer's instructions.

2.2.7 High efficiency transformation of the bait-containing yeast cells with the *Arabidopsis* cDNA library

Bait-containing yeast cells were inoculated in 70 ml Glu-L+M medium, and incubated at 24°C with shaking at 200 rpm for 72 hours. Cells were pelleted by centrifugation for 5 min at 1,000 x *g* in a 50 ml falcon tube. Cells were resuspended in 70 ml 2X YPD in a new flask and incubated at 24°C with shaking at 200 rpm for further 6 hours. Cells were pelleted as above, and resuspended in 10 ml LiSORB (100 mM LiAc, 1 M Sorbitol in 10 mM Tris-HCl pH 8.0, 1 mM EDTA-TE). 15 ml of LiSORB was added and mixed well. Cells were pelleted, and resuspended in LiSORB, and pelleted again as above. Cells were then resuspended in 2 ml LiSORB and rotated at room temperature for 30 min. ssDNA was boiled for 10 min and immediately cooled on ice for at least 5 min. 2 µg of the cDNA library was mixed together with 10 µl preboiled ssDNA in a 1.5 ml centrifuge tube. 180 µl of the above cell suspension was then added to the tube and mixed by vortexing. 1 ml of LiPEG (40% PEG_{3350} in 100 mM LiAc containing TE) was added and mixed by vortexing. Cells were incubated at room temperature with shaking for 30 min. 100 µl of DMSO (Sigma D2650) was added and mixed by vortexing. Cells were heat shocked for 10 min at 42°C and centrifuged at 3,000 x *g* for 1 min. The supernatant was discarded with a blue tip. After a short spin, the remaining LiPEG solution was removed completely with a yellow tip. The cells were resuspended in 150 µl of 1 M Sorbitol, plated on Glu-L-U+M plate, and incubated at 24°C for 5-7 days. 20 µl and 40 µl of 1:10 dilutions were plated out at the same time to facilitate calculation of transformation efficiency. A control plate was incubated at 36°C.

2.2.8 RRS screening

After 5-7 days of incubation, colonies were replica plated onto Gal-L-U-M plate and incubated at 36°C for 5-7 days. Colonies with sufficient growth were streaked onto Glu-L-U+M plate and incubated at 24°C for 3 days. The cells

were then replica plated onto four different plates: Gal-L-U-M, Glu-L-U-M, Gal-L-U+4M, YPD and incubated at 36°C for 3 days (Figure 2.8). Colonies that grew on Gal-L-U-M but not on the other three plates were streaked onto Glu-L-U+M plate, and incubated at 24°C for 3 days. The colonies were replica plated again onto the four plates described above. Colonies grew on Gal-L-U-M plate but not on the other three plates were streaked onto Glu-L-U+M plate and grown at 24°C for 3 days before stored at 4°C.

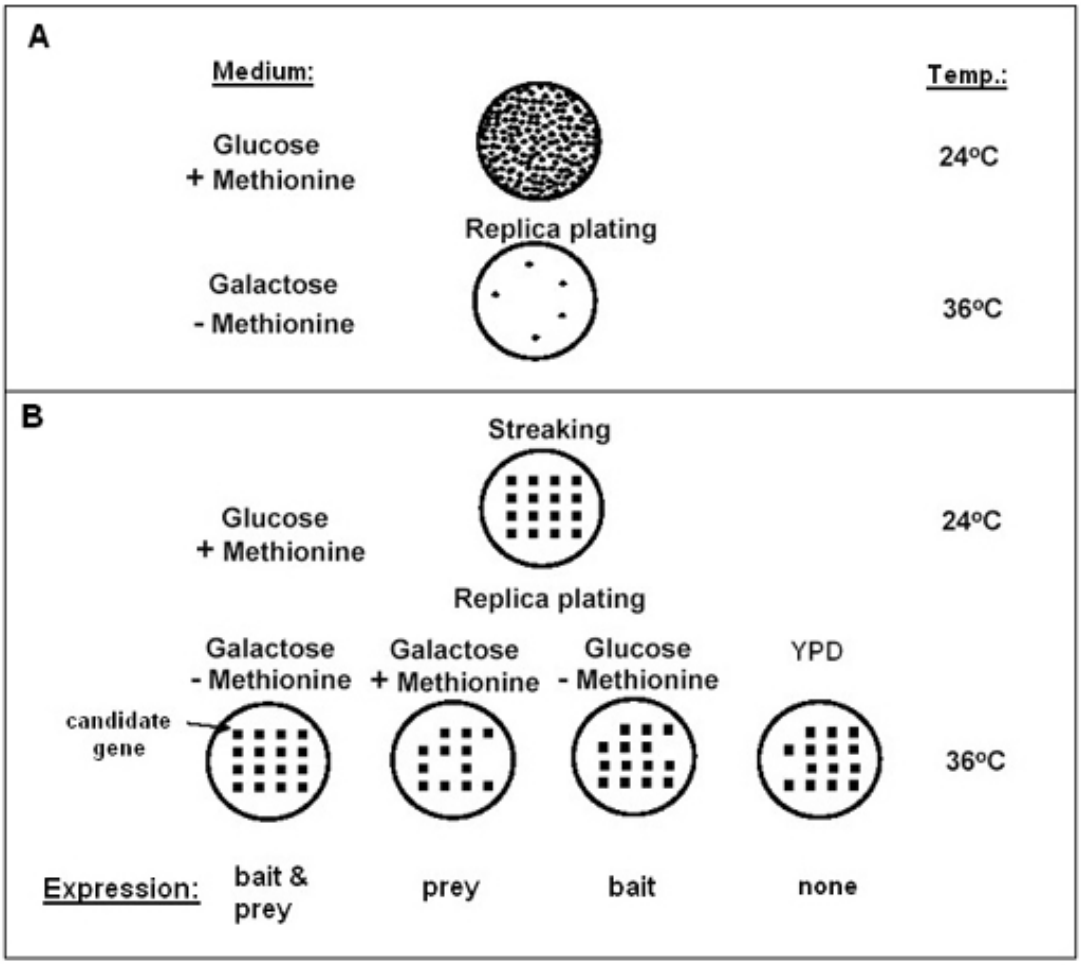


Figure 2.8 Flow chart of two-step GCR1 RRS selection screen. (A) Large-scale interaction screen identifies putative interactors. (B) Potential interactors are restreaked and replica plated onto 4 media to identify interactions that require expression of both bait and prey (adapted from Broder *et al*., 1998)

2.2.9 Yeast colony PCR

About 0.05 cm^2 yeast cells were harvested by scraping off the plate using an autoclaved toothpick. 6.5 µl lyticase mix (0.5 µl lyticase + 6 µl lysis solution (50 mM Tris-HCl pH7.5, 1.2 M sorbitol, 10 mM EDTA and 10 mM β-mecaptoethanol) was added and mixed well to resuspend cells. After cells were incubated at 37°C for 0.5 hour or overnight, 70 µl ddH_2O was added. The solution was boiled for 10 min, and cooled on ice. 1 µl of the solution was used as DNA template in a 10 µl PCR reaction (Table 2.10a and 2.10b).

2.2.10 *BLAST* searches

Sequence homology searches were performed using the BLAST algorithm available on The *Arabidopsis* Information Resource (TAIR) server (http://www.arabidopsis.org/wublast/index2.jsp). The BLAST program used in the search was BLASTN: NT query to NT db, and the dataset used was AGI Transcripts (-introns, +UTRs) (DNA). Default settings were used.

2.2.11 Dot-blot

2.2.11.1 Membrane preparation

Yeast samples were prepared the same as for yeast colony PCR (section 2.2.9). DNA samples were boiled for 10 min and immediately cooled on ice before loading to membrane. Nylon membrane (neutral, Sigma) was rinsed in 6X SSC for 10 min, and 25 µl of each denatured yeast DNA sample was loaded on it by hand using a pipette. Membrane was allowed to dry before it was rinsed in denaturating solution (0.5 M NaOH, 1.5 M NaCl) for 5 min, and in neutralizing solution (0.5 M Tris, 1.5 M NaCl, pH7.5) for 5 min. DNA was fixed to the membrane by UV crosslinking at 120 mj twice using the CL-1000 UV crosslinker (UVP). 1:1000 dilutions of purified PCR products of identified interactors were used as positive controls. 1:1000 dilution of purified PCR product of GCR1 was used as negative control.

2.2.11.2 Probe preparation

Digoxigenin-dUTP labelled probes were synthesised using the PCR DIG Probe Synthesis Kit (Roche), as described in table 2.17 and 2.18. The primers used for probe synthesis are listed in appendix 3. DIG-labelled PCR products were examined by electrophoresis, and purified using the High Pure Product Purification Kit (Roche). Probes were denatured by boiling for 5 min and rapidly cooled on ice prior to use. Denatured DIG-labelled DNA probes were added to pre-heated DIG Easy Hyb (3.5 ml/100 cm^2 filter, at least 6 ml per roller bottle) and mixed well. When probes were reused, probe-DIG Easy Hyb mixture was denatured before use by heating to 68°C for 10 min. For the lowest possible background, the probe-DIG Easy Hyb mixture was filtered through a 0.45 µm cellulose acetate filter before adding to the membrane.

Table 2.17 PCR amplification of digoxigenin-dUTP labelled probes

Reagents	Vol (ul)
10x PCR buffer with $MgCl_2$	5
PCR DIG labelling mix	5
Enzyme mix	0.75
Forward primer (10 uM)	2.5
Reverse primer (10 uM)	2.5
ddH_2O	29.25
DNA template	5

	Steps	Temp (°C)	Time
1	Denaturation	95	2'
2	Denaturation	95	30"
3	Annealing	55	30"
4	Extension	72	2' 30"
5	Repeat (2-4)	-	34 cycles
6	Final polishing	72	7'

Table 2.18 Probes used in dot-blots

Probe	Primers used to amplify the probe	l* (bp)	T_{hyb} (°C)
TRX3	TRX3 Beg + TRX3 R1	357	41.92-46.92
TRX4	TRX4 Beg + TRX4 R1	360	40.78-45.78
CL12C	CL12C F1 + CL12C R1	301	41.90-46.90
PRL27aC	PRL27aC F1 + PRL27aC R1	313	42.82-47.82
P450	P450 F1 + P450 ER1	126	38.28-43.28
PDF2.2	PDF2.2 Beg + PDF2.2 ER1	240	43.44-48.33
ADPGlu PPase	ADPglcP F1 + ADPglcP ER1	147	38.20-43.20
C2dcp	C2dcp F1 + C2dcp R1	189	38.79-43.79
MLP168	MLPR F1 + MLPR ER1	206	41.41-46.41
Glyrich	Glyrich F1 + Glyrich R1	142	44.55-49.55
CL12A	CL12A Beg + CL12C R1	301	41.90-46.90
PRMT	MT F2 + MT R1	124	36.83-41.83
ASK1	ASK1 F1 + ASK1 R1	177	39.27-44.27
AHB2	HB2 F1 + HB2 R1	218	40.32-45.32
PAG1	PAG1 F1 + PAG1 R1	173	40.78-45.78
AGL42	AGL42 F1 + AGL42 R1	111	34.91-39.91
MLP43	BetVI KpnIbeg + BetVI R1	456	41.59-46.59
Thionin	Thionin Beg + Thionin R1	405	41.77-46.77
QPRTase	QPRTase F1 + QPRTase R1	261	41.37-46.37
LRR RLK	LRR RLK F1 + LRR RLK R2	147	38.03-43.03
SRO5	SRO5 F1 + SRO5 R1	182	36.69-41.69
ZAC	ZAC F1 + ZAC R1	148	37.17-42.17
40SRPS2	40SRPS2 Beg + 40SRPS2 R1	855	43.80-48.80
TRX2	TRX2 F1 + TRX2 R1	357	41.18-46.18
zf-DHHC1	ZF F1 + ZF R2	642	42.34- 47.34
ExPro	ExPro F1 + ExPro R1	177	39.47-44.47

* l=length of hybrid in base pairs. T_{hyb} is determined according to the following equation: $T_{hyb} = T_m$ – 20 to 25℃. T_m = 49.28+0.41(%G+C)-(600/l). GC% is calculated a using online program:
http://www.artsci.wustl.edu/~twest/molbio/gccontent.php

2.2.11.3 Hybridisation and immunological detection

DIG High Prime DNA labeling and Detection Starter Kit II (Roche) was used for dot-blot hybridisation and immunological detection. Pre-hybridisation and hybridisation were performed in a roller bottle with gentle agitation in a hybridisation oven (Bachofer, Germany). Membrane was pre-hybridized at T_{hyb} (43°C) in 10 ml/100 cm^2 of DIG Easy Hyb hybridisation buffer (at least 20 ml per roller bottle) for 30 min, and hybridized at T_{hyb} (43°C) overnight in 3.5 ml/100 cm^2 of fresh hybridisation buffer containing denatured DIG-labelled probe (2-4 µl/ml hybridisation buffer).

Membrane was washed 2x 5 min in 2x SSC (0.3 M NaCl, 0.03 M Nacitrate, pH7.0) plus 0.1% SDS at room temperature and 2x 15 min in 0.5x SSC, 0.1% SDS at 68°C under constant agitation. Immunological detection was carried out as follows: incubated for 30 min in 100 ml/100 cm^2 membrane of 1x blocking solution; incubated for 30 min in 20 ml/100 cm^2 membrane of antibody solution (75 mU/ml in blocking solution); washed 2x 15 min in 100 ml/100 cm^2 membrane washing buffer (0.1 M Maleic acid, 0.15 M NaCl, pH7.5; 0.3% v/v Tween 20); equilibrated 5 min in 20 ml/100 cm^2 membrane detection buffer (0.1 M Tris-HCl, 0.1 M NaCl, pH9.5). Membrane was treated with chemiluminiscent CSPD ready-to-use and the luminescence was detected by exposing to high performance autoradiography film (Hyperfilm™ MP, Amersham Biosciences) for 2 hours. Films were developed using the OPTIMAX X-Ray Film Processor (PROTEC, Germany) and scanned on a flat bed scanner filled with a transparency adaptor and images preserved using Adobe Photoshop.

2.2.11.4 Stripping and reprobing

DIG-labelled probe was removed from membrane filter by rinsing in ddH_2O for 5 min, washing in 0.2 M NaOH, 0.1% (w/v) SDS for 15 min at 37°C twice. After rinsing thoroughly with 2x SSC, membrane was hybridised to another probe. Stripped membrane filter was stored in 2x SSC at 4°C.

2.2.12 Plasmid rescue

2 cm^2 of yeast cells were collected using a flat headed toothpick and resuspended in 60 µl of lysis solution plus 5 µl of lyticase. The cells were incubated at 37ºC for 2 hours or overnight and collected by centrifugation at 1,500 x *g* for 5 min in a 1.5 ml centrifuge tube. The supernatant was removed and the pellet was used to isolate plasmids using the Wizard *Plus* SV Miniprep DNA Purification System (Promega). 80 µl of the isolated plasmid DNA was freeze-dried to 5 µl and transformed with 25 µl of competent *E. coli* cells.

Two sets of primers were used in the colony PCR for each colony. The pMetF and LP2R primers were used to identify the bait containing plasmid. The T7 and a gene (prey) specific reverse primer were used to identify prey containing plasmid. The primer sequences are listed in appendix 3. The colony containing the prey plasmid only was patched out on a LB-amp plate and the plasmid was isolated using the Wizard *Plus* SV Miniprep DNA Purification System (Promega).

2.2.13 Protein expression in *E.coli*

2.2.13.1 Protein expression induction by IPTG

A single *E.coli* colony was inoculated into 5 ml LB containing appropriate antibiotics in a 50 ml tube and incubated at 37°C with shaking at 250 rpm overnight. 0.5 ml of the overnight cell culture was inoculated into 9.5 ml LB containing appropriate antibiotics and incubated at 37°C with shaking at 250 rpm. Readings of A600nm were taken to monitor the growth of the cells. When A600nm reached 0.4 – 0.6 (about 1.5 – 2 hours), 2 ml of cells at an OD of 0.5 were harvested by centrifugation at 4,000 x *g* for 5 min. The supernatant was discarded and the cells were resuspended in 50 µl HEPES buffer (50 mM HEPES, 0.2 M NaCl, pH8.0). This uninduced sample was kept

at -20°C until further use. IPTG (0.5 mM, final concentration) was added to the rest of the culture and cells were incubated at 37°C or 25°C with shaking at 250 rpm until A600nm reached 0.8 – 1.0 (about 3 hours) or at 18°C overnight. 1 ml of cells at an OD of 1.0 were harvested by centrifugation at 4,000 x *g* for 5 min. The supernatant was discarded and the cells were resuspended in 50 µl HEPES buffer. This induced sample was kept at -20°C until further use.

2.2.13.2 Sample preparation of total cell fraction for SDS-PAGE

Expression of the target gene was assessed quickly by analysis of the total cell fraction on a SDS-PAGE followed by Coomassie blue staining. To obtain the total cell protein, 50 µl of 2x sample buffer (0.12 M Tris, 10% v/v Glycerol, 4% SDS, 0.005% Bromophenol blue, 10% v/v β-mercaptoethanol) was added to each of the 50 µl uninduced or induced sample. The sample was heated at 95°C for 5 min and then sonicated for 3x 3sec on ice using a probe sonicator (MSE) at an 18 micron amplitude setting. The sample was centrifuged for 5 min at 16,000 x *g* and 10 µl was use for SDS-PAGE.

2.2.13.3 Protein solubility check

10 ml of induced culture was centrifuged at 4,000 x *g* at 4°C for 10 min. The supernatant was discarded and the cells were resuspended in 500 µl HEPES buffer containing 1% (final concentration) Triton X-100. The sample was frozen on dry ice and thawed in a room temperature water bath. The cycle was repeated three times for 10 min intervals. The sample was sonicated 6x 10 sec on ice and then centrifuged at 16,000 x *g* for 10 min. The supernatant (soluble fraction) was saved in a new tube and the pellet (insoluble fraction) was resusupended in 500 µl HEPES buffer. 10 µl each of soluble and insoluble fraction was mixed with 10 µl 2x sample buffer and heated at 95°C for 5 min. 10 µl of the mixture was used for SDS-PAGE.

2.2.13.4 Preparation of cleared cell lysate

Protein expression was induced by addition of IPTG as described in 2.2.13.1. Cells were harvested by centrifugation at 4,000 x *g* for 10 min at 4°C. Each gram of cells was resuspended in 5 ml of lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH8.0). The suspension was lysed by three cycles of freeze – thaw and 6x 10 sec of sonication as described in 2.2.13.4. It was then centrifuged at 4,000 x *g* for 10 min at 4°C. The supernatant was transferred to a new tube and stored at -80°C.

2.2.14 Protein purification

100 µl of His-select affinity beads (Sigma) were centrifuged at 5,000 x *g* for 30 sec. The supernatant was removed and the beads were washed with 300 µl of equilibration buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH8.0). 200 µl of cell lysate was incubated with the beads at room temperature for 30 min. The beads were washed five times with 1 ml washing buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH8.0) and eluted with 100 µl of elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 400 mM imidazole, pH8.0).

2.2.15 Pull-down assay

800 µl of cell lysate containing the bait protein was incubated with 100 µl of His-affinity beads at room temperature for 30 min to capture the His-tagged bait protein. The beads were then washed 5x 1 ml of washing buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH8.0) to remove unspecific binding proteins. 50 µl of cell lysate containing the prey protein was incubated with 2 mM DTT at 25°C for 3 hours before being added to the above bait containing beads. The mixture was incubated at 4°C for 16 hours. The beads were washed 5 times in the same manner and finally resuspended in 100 µl of washing buffer. 50 µl of the beads were mixed with 50 µl of sample buffer and boiled for 5 min, and the other 50 µl of the beads were eluted with 100 µl of elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 400 mM imidazole, pH8.0).

2.2.16 Growth of plant material

To obtain mature plant material, seeds were sown into pots containing pre-wetted Levington F2 compost (Levington, Ipswich, UK) supplemented with 2gL^{-1} Osmocote, a slow release fertilizer and 20% v/v Perlite (Levington). A top drench of the insecticide 0.02% w/v Intercept 70 WG solution was applied (100 ml per liter compost) before planting the seeds. After stratification at 4°C for 3 days, they were transferred to a controlled environment glasshouse. They were grown at 23°C during the day and 18°C as the night temperature, under long day conditions (16 hour photoperiod, supplemented by light when necessary).

2.2.17 Transformation of *Agrobacterium tumefaciens* with plasmids

30 ng of plasmid was incubated with 30 µl competent *Agrobacterium tumefaciens* GV3101 on ice. The mixture was transferred to a prechilled electroporation cuvette (Gene Pulser Cuvette, 0.1cm electrode gap, Bio-Rad). The electroporation was carried out using the MicroPulser electroporator (Bio-Rad) using the Agr programme (Voltage: 2.2 kV, Time constant: 5 msec). Immediately after electroporation, 1 ml of 2xYT medium was added to the cuvette, and the bacterial suspension was transferred to a 1.5 ml eppendorf tube and incubated at 28°C for 2 – 3 hours with gentle agitation. 100 µl of the cells were spread onto a 2xYT plate containing the appropriate antibiotics (rifampicin 100 µg/µl, gentamycin 25 µg/µl, ampicillin 100 µg/µl, kanamycin 50 µg/µl).

2.2.18 Infiltration of *Nicotiana benthamiana* with *Agrobacterium tumefaciens*

Agrobacterium tumefaciens cells were grown in 10 ml 2xYT medium containing kanamycin (50 µg/µl) and silencing suppressor p19 (equivalent to

10 µl of cells at an A600 of 1.0 OD) at 28°C with shaking of 225 rpm for 24 hours. The culture was centrifuged and resuspended in 20 ml infiltration buffer (50 mM $MgCl_2$, 100 µM acetosyringone). The A600 of the suspension was measured and adjusted to 1.0. The cells were incubated at room temperature with gentle shaking for 3 hours before infiltration. The *Agrobacterium tumefaciens* harboring different plasmids were mixed in equal volumes and infiltrated (~ 3 ml/leaf) into tobacco leaves from the underside of the leaf using a 5 ml syringe without a needle. After infiltration, the plants were further grown in the greenhouse. Leaf discs of infiltrated plants were taken everyday up to day 6. The harvested plant material was frozen at -80°C for storage. 1 cm^2 of leaf disc was grinded in 50 µl SDS-PAGE sample buffer, boiled for 5 min and centrifuge at 16,000 x *g* for 5 min. 15 µl of the supernatant was used in the Western blot to check protein expression.

2.2.19 Amino acid multiple sequence alignment

The amino acid multiple sequence alignment was performed using the online programme ClustalW. Default settings were used.
http://www.ebi.ac.uk/Tools/clustalw2/index.html

Chapter 3 Making bait constructs for use in the RRS screening

3.1 Introduction

The yeast two-hybrid system is the most commonly used method for large-scale, high-throughput identification of potential protein-protein interactions (Walhout *et al*., 2000). In this system, the protein of interest is normally used as the bait to fish out the prey protein(s) that it interacts with from a pool of proteins. In the RRS (section 1.2.4.2), the coupling of the bait to the prey brings mRas to the cell membrane where it can interact with other signalling molecules that lead to cell growth at restrictive temperature. This also means that the TM regions of membrane proteins are not suitable for use as baits in the RRS, as they will locate the mRas to the membrane hence autoactivate the system. A large number of studies have suggested that the intracellular loops of GPCRs are the key regions for G-protein coupling (Wong 2003). For example, deletions at the 2nd and 3rd intracellular loops have resulted in GPCRs that no longer couple to G-proteins while retaining their ligand-binding conformation (Takhar *et al*, 1996; Chicchi, 1997). The observation that the GCR1-GPA1 interaction in *Arabidopsis* depends on the intracellular domains of GCR1 (Pandey and Assmann, 2004) indicates that the intracellular domains of GCR1 may be the most important regions that interact with its downstream targets, and that they can be used as baits in the RRS screening for GCR1 interactors.

GPCRs are integral membrane proteins with seven transmembrane regions. Membrane proteins form approximately 20% of most genomes (Taylor *et al*., 2006). They are insoluble in aqueous solution and nonpolar solvent, and are difficult to crystallize because of the membrane lipids or detergents that are necessarily bound to their nonpolar surfaces (Creighton, 1993). It is inherently difficult to obtain information about their 3-dimensional (3D) structure using

the main experimental methods for structure determination, such as X-ray crystallography and multidimensional nuclear magnetic resonance (NMR) spectroscopy. To date, the 3D structure of GCR1 has not been determined.

However, information about the structure of membrane proteins can be gained by studying their membrane topography. The basis of the cellular membrane structure is the amphiphilic nature of the lipid molecules. The polar, usually charged head groups of the lipids contact with water and the nonpolar tails aggregate side by side to avoid contact with water. Therefore, the transmembrane (TM) regions, which span the membrane, are predominantly hydrophobic and normally contain 17-25 amino acids to provide sufficient length to cross the membrane (Taylor *et al*., 2006). These features make it possible to predict the TM regions of a membrane protein from its amino acid sequence alone.

In fact, the TM regions of proteins of unknown tertiary structure have been identified simply by searching their primary structures for hydrophobic stretches of amino acid residues that are similar to those in the known membrane proteins (Creighton, 1993). Fortunately, computer programs have been developed, which have made the protein topography prediction process fast and relatively easy. The inclusion of the 'positive-inside rule' whereby non-membrane regions inside the cell have more positively charged residues, such as arginine and lysine, than the regions outside (von Heijne 1986), and the use of hidden Markov models (HMMs) and other machine-learning techniques, had led to a significant improvement in the predictions using these computer programs (Elofsson and Heijne, 2007).

This project aimed to identify GCR1 interacting proteins using the RRS, therefore the first objective is to accurately predict GCR1 topography so that its intracellular regions can be identified and used as baits in the subsequent library screening for GCR1 interactors. Described below are the three computer programs that I have used to predict GCR1 topography.

3.1.1 Membrane protein topography prediction programs

Several transmembrane prediction websites are available, so in order to most accurately determine the 7TM regions of GCR1 it was analysed using the three most widely used prediction programs.

3.1.2.1 TMpred – Transmembrane Prediction

The TMpred program makes a prediction of membrane-spanning regions and their orientation. The algorithm is based on the statistical analysis of TMbase, a database of transmembrane proteins and their helical membrane-spanning domains. TMbase is based primarily on SwissProt accessions, but contains sequence information from other sources as well. The prediction is made using a combination of several weight-matrices for scoring. The influences of neighbouring residues, membrane protein classification, taxonomic classification and segment orientation on these positional preferences are taken into account during the analyses (TMbase, 2008).

3.1.2.2 TMHMM – Tied-mixture Hidden Markov Modeling

Hidden Markov model (HMM), a probabilistic framework, has been widely used in computational biology to model gene structure and the statistical structure of genomes, to generate profiles for protein families and to align sequences (Sonnhammer, *et al*., 1998; Tusnady and Simon, 1998; Kahsay *et al*., 2005). A special application of the sequence alignment is in protein topography prediction using secondary structure sequences (Tusnady and Simon, 1998). The basic principle is to define a set of states, each corresponding to a region or specific site in the proteins being modelled (Sonnhammer, *et al*., 1998).

In TMHMM, the model comprises seven sets of states, with each set corresponding to a type of region in the protein sequence: one for the transmembrane helix core, two for helix caps on either side, one for the loop

on the cytoplasmic side, one each for the short and long loops on the non-cytoplasmic side that correspond to two different membrane insertion mechanisms, and one for 'globular domains' in the middle of each loop. Each set of states has an associated probability distribution over the 20 amino acids characterising the compositional bias in the corresponding regions. The states are connected to each other in a biologically reasonable way. For example, the state for inside loop is connected to itself and to the transmembrane helix state, because a loop may be longer than 1, and after a loop a helix begins. The model defines for each residue and connecting them in a cycle. The path of a protein sequence through the states with the highest probability predicts the topography (Sonnhammer, *et al*., 1998; Kahsay *et al*., 2005).

3.1.2.3 TMAP (single)

TMAP predict membrane protein topography from multiply aligned amino acid sequences after determination of the membrane-spanning segments. The prediction technique relies on residue compositional difference in the protein segments exposed at each side of the membrane. Intra/extracellular ratios are calculated for the residue types Asn(N), Asp(D), Gly(G), Phe(F), Pro(P), Trp(W), Tyr(Y) and Val(V), most found on the extracellular side, and for Ala (A), Arg(R), Cys(C) and Lys(K), mostly occurring on the intracellular side. The consensus over these 12 residue distribution is used for sidedness prediction.

3.1.2 Research objectives and experimental approach

This part of this project aimed to: (1) Predict GCR1 transmembrane topography using the three computer programs mentioned above then determine the regions for use as baits in the RRS; (2) Clone these regions into the yeast expression vector to make the bait constructs and transform them into yeast; (3) Check baits protein expression in yeast; (4) Test whether

the baits are suitable for use in the RRS screening by checking whether they autoactivate the system or not.

3.2 Results

3.2.1 GCR1 transmembrane topography prediction and choice of bait sequences

The transmembrane topography of GCR1 was predicted using three transmembrane protein topography computer programs: TMpred, TMHMM and TMAP. All three programs predict GCR1 to be an integral membrane protein, which has an extracellular Nter, seven TM regions linked by three intracellular loops (i1, i2, i3) and three extracellular loops (e1, e2, e3), and a cytoplasmic Cter (Figure3.1).

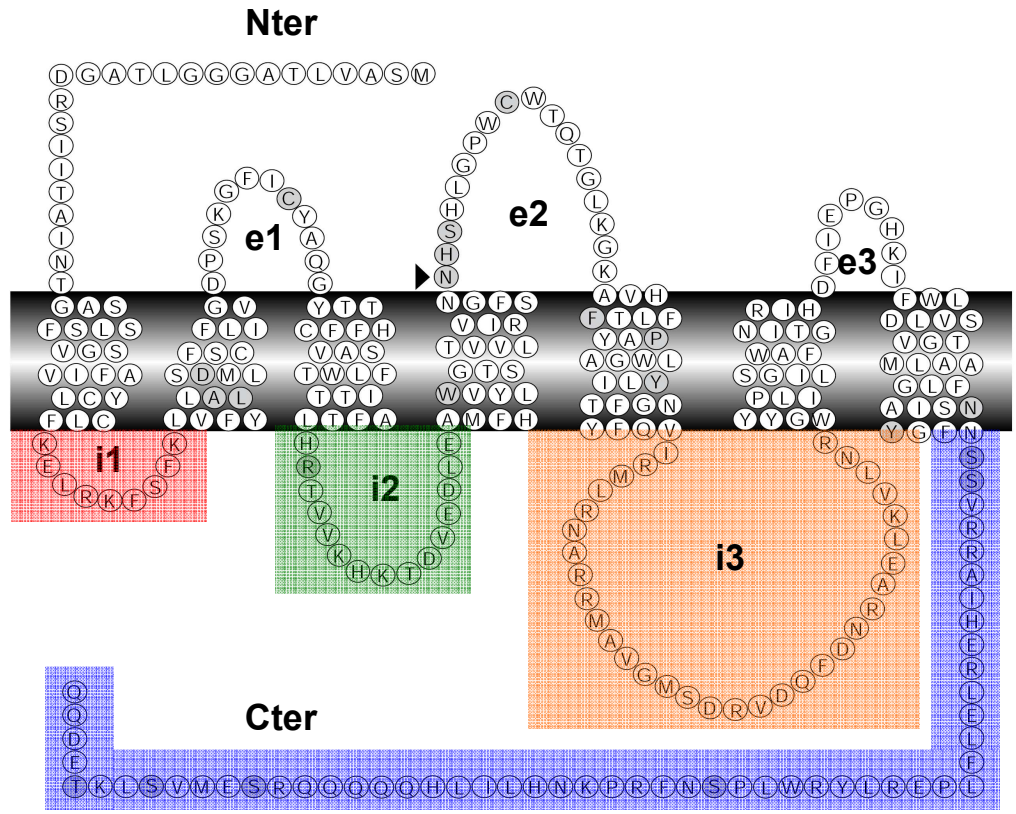


Figure 3.1 GCR1 topography (adapted from Plakidou-Dymock *et al*., 1998)

The boundaries of the intracellular regions i1, i2, i3 and Cter predicted by the three programs were slightly different as shown in table 3.1. Since the intracellular regions would be used as the bait sequences, and bait binding to the plasma membrane will result in autoactivation of the RRS, the chemical properties of the boundary region amino acids were carefully analysed. The hydrophobic amino acids at the boundaries which could potentially bind to the membrane were excluded from the bait sequences. The choice of the baits is illustrated in table 3.1 and figure 3.2. Because the i1 and i2 loops were so short it was decided to fuse them together into a single bait for initial RRS screens.

Table 3.1 Predicted positions of GCR1 intracellular regions

GCR1 Regions (Baits)	**Positions predicted by different programs**					**Choice of positions of GCR1 regions**
	TMpred			TMHMM	TMAP	
	17aa*	19aa*	21aa*			
i1	45-53	45-53	45-53	43-53	41-50	45-53
i2	106-120	106-120	106-120	107-118	103-119	106-120
i3	186-217	186-217	186-217	184-217	187-218	184-217
Cter	273-326	273-326	273-326	270-326	274-326	270-326

*Three TMpred predictions were made using 17aa, 19aa and 21aa as the minimum length of hydrophobic helix respectively. Default settings were used for TMHMM and TMAP predictions. Numbers indicate the amino acids residue at the start and end of each intracellular region. The final column (shaded) shows the choice of the positions of GCR1 intracellular regions (baits).

```
  1  MSAVLTAGGG LTAGDRSIIT AINTGASSLS FVGSAFIVLC YCLFKELRKF
                                                        i1
 51  SFKLVFYLAL SDMLCSFFLI VGDPSKGFIC YAQGYTTHFF CVASFLWTTT

101  IAFTLHRTVV KHKTDVEDLE AMFHLYVWGT SLVVTVIRSF GNNHSHLGPW
              i2
151  CWTQTGLKGK AVHFLTFYAP LWGAILYNGF TYFQVIRMLR NARRMAVGMS
                                                  i3
201  DRVDQFDNRA ELKVLNRWGY YPLILIGSWA FGTINRIHDF IEPGHKIFWL
```

251 SVLDVGTAAL MGLFNSIAYG FNSSVRRAIH ERLELFLPER LYRWLPSNFR

Cter

301 PKNHLILHQQ QQQRSEMVSL KTEDQQ*

Figure 3.2 Choice of GCR1 intracellular regions. GCR1 amino acid sequence is shown above. The intracellular regions i1, i2, i3 and Cter are underlined.

3.2.2 Making the pMetRas-i1-i2, pMetRas-i3, pMetRas-Cter bait constructs

The inserts i1-i2, i3, Cter were amplified by PCR (Figure 3.3) using the KOD DNA polymerase (section 2.2.2.1). Two rounds of PCR were performed in order to make the i1-i2 insert. In the first round PCR, $i1^+$ (from GCR1 start codon to the end of i1) and $i2^+$ (from the beginning of i2 to the end of i3) were amplified separately using the *GCR1* DNA as template, with the GCR1bgn + LP1R and the LP2F + LP3RXmaI primer pairs respectively. In the second round PCR, i1 and i2 were joined together using the 1:100 diluted first round $i1^+$ and $i2^+$ PCR products as templates, and using the LP1FXmaI primer + LP2RXmaI primer pair (Table 2.5). The i3 and Cter inserts were amplified using the *GCR1* DNA as template, and using the LP3FXmaI + LP3RXmaI and the CterFXmaI + CterRXmaI primer pairs respectively (Table 2.5). The predicted size of each insert is i1-i2: 99 bp, i3: 132 bp, Cter: 201 bp and the products obtained were consistent with this.

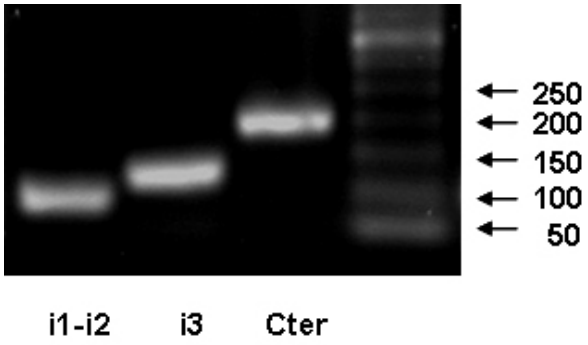


Figure 3.3 PCR amplification of the inserts. The three inserts, GCR1 i1-i2, i3, and Cter were PCR amplified. The predicted size of each insert is: i1-i2: 99 bp, i3: 132 bp, Cter: 201 bp.

PCR products were digested with *XmaI* (New England Biolabs, NEB) and ligated into *XmaI* digested bait expression vector pMetRas as described in sections 2.2.2.1-2.2.2.3. Recombinants were checked for insert size and orientation by colony PCR with a vector-specific primer pMetF as forward primer, and a bait-speicifc primer such as LP1RXmaI as reverse primer (Figure 3.4). Positive clones were identified and the plasmids were extracted and sequenced (data not shown). Sequencing results confirmed that all inserts were successfully cloned in frame to the N-terminus of Myc which is translational fused to the N-terminus of Ras in pMetRas (*XmaI* site), which generated the pMetRas-i1-i2, pMetRas-i3, and pMetRas-Cter constructs.

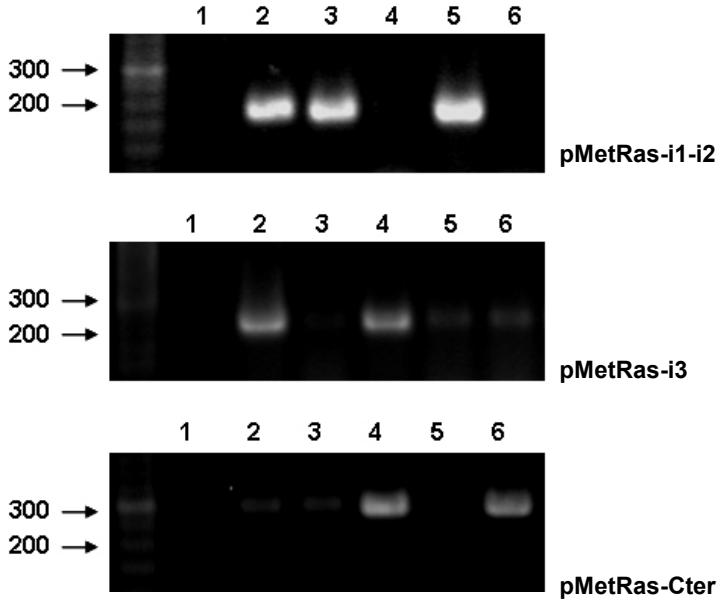


Figure 3.4 Colony PCR. Colony PCR was conducted with a vector-specific primer as forward primer, and a bait-speicifc primer as reverse primer to screen for positive clones. Positive clones were found for i1-i2 (lane 2, 3 and 5), i3 (lane 2 and 4) and Cter (lane 4 and 6).

3.2.3 Detecting the expression of the i1-i2, i3 and Cter baits

The pMetRas-i1-i2, pMetRas-i3 and pMetRas-Cter constructs, and the empty bait vector pMetRas which was used as a control, were transformed into the temperature sensitive yeast strain *cdc25-2*. Total proteins were extracted from the yeast cells that were grown in Glu-L-M media for bait expression and

Glu-L+4M media for suppression of the bait. Because the bait vector contains a Myc epitope tag, the expression of the bait fusion proteins was checked by Western blot using the anti-Myc antibody. In the absence of methionine (-methionine), polypeptides of the expected size for the i1-i2-Myc-Ras and i3-Myc-Ras bait fusions and the Myc-Ras control were expressed at a high level whereas polypeptide for the Cter-Myc-Ras fusion was expressed at a very low level. In the presence of methionine (+ methionine), very low level of expression was detected for the i1-i2-Myc-Ras fusion and the Myc-Ras control, and no expression was detected for the i3-Myc-Ras and Cter-Myc-Ras fusions (Figure 3.5). The Western blot was repeated twice more and the same results were obtained.

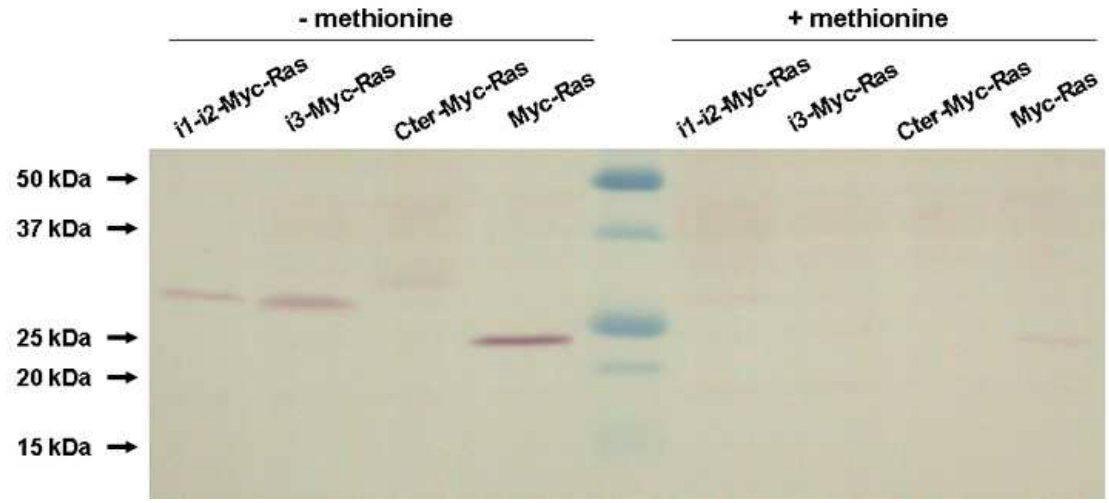


Figure 3.5 Western blot detection of bait-Myc-Ras fusion protein expression. Total proteins were extracted from yeast cells that were grown in Glu-L-M media (-methionine) for bait expression and Glu-L+4M media (+methionine) for bait suppression. The expression of the bait fusion proteins was checked by Western blot using the anti-Myc antibody. The expected size of each fusion protein is: i1-i2-Myc-Ras: 26 kDa, i3-Myc-Ras: 27.1kDa, Cter-Myc-Ras: 30 kDa, and Myc-Ras (control): 23.05 kDa.

3.2.4 Bait autoactivation test and selection of suitable baits

Four yeast colonies for each of the three baits (i1-i2, i3 and Cter) were randomly selected for the temperature sensitivity test. They were streaked onto Glu-L+M plates and grown at 24°C for 3 days. These bait-containing

yeast cells were then replica plated onto two Glu-L+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as control. No bait-containing yeast cells grew at 36°C, but all grew at 24°C (Figure 3.6). Therefore these bait-containing yeast cells were temperature sensitive and could be used in the following experiments.

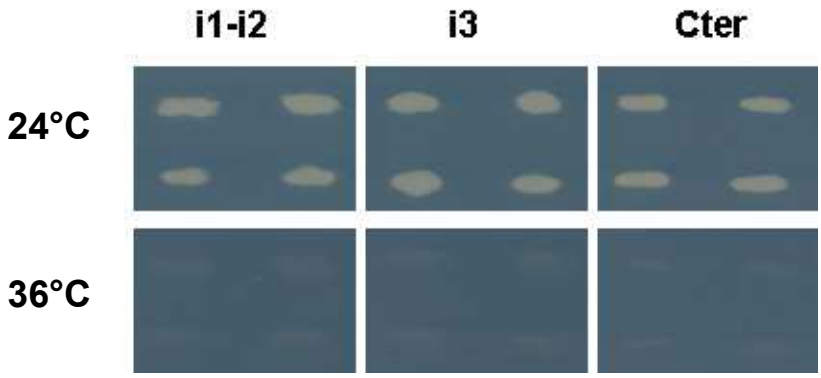


Figure 3.6 Bait temperature sensitivity test. The bait-containing (i1-i2, i3 or Cter) yeast cells were streaked onto Glu-L+M plate and grown at 24°C for 3 days. The colonies were then replica plated onto two Glu-L+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as control. Photos show growth after this 3 day period.

It is vital to work only with baits that do not autoactivate the RRS. To check for autoactivation, yeast cells expressing bait with no prey were replica plated onto two Glu-L-M plates, one was grown at 36°C for 5 days, and the other at 24°C for 5 days as control. All yeast cells grew at 24°C as expected. No yeast cells containing the pMetRas-i1-i2 construct grew at 36°C, but all cells containing the pMetRas-i3 or pMetRas-Cter constructs grew at 36°C (Figure 3.7). The result demonstrated that the i1-i2 bait did not autoactivate the RRS, whereas the i3 and Cter baits did autoactivate the RRS in the absence of preys. So the i1-i2 bait was suitable for use in the RRS, but the i3 and Cter baits were not suitable baits for use in the RRS. Therefore, only the i1-i2 bait construct was used in the subsequent library screening using the RRS.

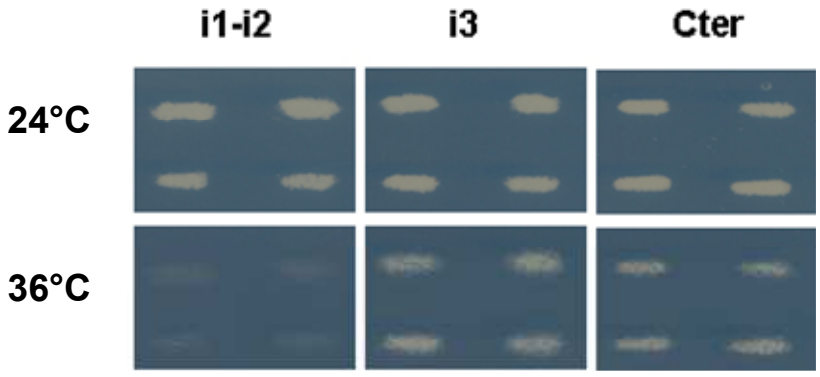


Figure 3.7 Bait autoactivation test. To check whether the baits autoactivate the RRS or not, the bait-containing yeast cells were replica plated from Glu-L+M plate onto two Glu-L-M plates, one was grown at 36°C for 5 days, and the other at 24°C for 5 days as control. Photos show growth after this 5 day period.

3.3 Discussion

3.3.1 GCR1 transmembrane topography prediction

The availability of the crystal structures of rhodopsin and β2AR has confirmed the 7TM structure of GPCRs determined by hydrophobicity analysis on the primary sequences (Palczewski *et al*., 2000; Wong, 2003; Cherezov *et al*., 2007; Rasmussen *et al*., 2007; Rosebaum *et al*., 2007). The transmembrane spans are α-helical segments interconnected by alternating extracellular and intracellular loops bundled together to form an α-helical bundle-like structure (Palczewski *et al*., 2000). So far, the Protein Data Bank (PDB) holds only ~100 high – resolution structures of integral membrane proteins of the α-helix bundle type (von Heijne, 2006), representing less than 1% of the proteins of known structure (Chen and Rost, 2002; Elofsson and Heijne, 2007). Among these 100 membrane proteins of known structure, 14 are GPCRs. However, the crystal structure of GCR1 has not been obtained.

Topography prediction methods are convenient and less expensive ways of predicting protein structures than crystal structure determination. The accuracy of the topography prediction reached 70-80%, while the accuracy of the prediction of the transmembrane helices reached 90-95% (Tusnady and

Simon, 1998). Despite the lack of crystal structure information on most membrane proteins, many research groups have used various topography prediction methods to predict the topography of their protein of interest. For example, Devoto *et al*. (1999) used different programs to predict the topography of a MLO protein (a family of integral membrane proteins that are involved in plant defense and cell death). Membrane topography predicted by the TMHMM program suggested the existence of seven hydrophobic segments with the potential to form transmembrane helices. By using a combination of scanning *N*-glycosylation mutagenesis and Lep-MLO fusion proteins, they proved that MLO is membrane-anchored by 7TM helices with its Nter located extracellularly and Cter intracellularly.

Given that each computational prediction program has its strength and weakness, a sequence should be examined by several programs and all results should be taken into consideration. In this project the same orientation and number of TM domains were predicted for GCR1 when three different programs were employed. GCR1 was predicted to be a membrane protein with an extracellular Nter, seven TM domains linked by three extracellular and three intracellular loops, and an intracellular Cter. It is known that a slight change of conformation can alter efficacy and selectivity of G-protein activation and this enables the receptor-G-protein interaction to be fine-tuned by different extracellular ligands as well as by the intracellular environment (Wong 2003). For instance, it has been shown that a conserved set of residues on the cytoplasmic surface of rhodopsin, where G-protein activation occurs, likely undergo a conformational change upon photoactivation of the chromophore that leads to rhodopsin activation and signal transduction (Palczewski *et al*, 2000). This reveals that the boundaries of the intracellular domains may move up and down in relation to the membrane when the protein is in different conformations. It is hence difficult to predict precisely the boundaries of the intracellular region using the topography prediction programs. Although the programs employ different prediction method (section 3.1.1), the boundaries of each intracellular region predicted by the three

programs were highly similar, with only 0 – 4 amino acids differences (Table 3.1). Considering that bait binding to the plasma membrane will result in autoactivation of the RRS, we made fine adjustment at the boundaries of each intracellular region by excluding hydrophobic amino acids at the boundaries from the bait constructs in an attempt to reduce the probability that the bait would bind to the membrane and hence autoactivate the RRS. It is recognized that this could have eliminated regions of the GCR1 protein that may have a role in protein-protein interactions. While the i1-i2 bait was found to be suitable for the RRS, the i3 and Cter baits showed autoactivaiton of the system. This is discussed in the section below.

3.3.2 GCR1 intracellular regions as baits in the RRS

The intracellular regions of GCR1 were cloned separately into the bait expression vector for use in the RRS screens. Considering that both i1 and i2 are very short and may not function as good baits, they were joined together by a 2-step PCR (i1-i2, section 3.2.2) to be used as a single bait in the library screening. The strategy envisaged at this stage was to select preys interacting with this combined bait then determine which of the two regions they interact with in subsequent screening with individual i1 and i2 baits.

In the absence of methionine, the i1-i2-Myc-Ras and i3-Myc-Ras bait fusion proteins were expressed at a high level and the Cter-Myc-Ras fusion was expressed at a very low level (Figure 3.5). This confirmed that the bait expression can be induced by incubation in media lacking methionine. It also suggests that the induction level depends on individual baits. The addition of methionine in the media is supposed to suppress bait expression. In the presence of methionine, no expression was detected for the i3-Myc-Ras and Cter-Myc-Ras bait fusion proteins, but a very low level of expression was detected for the i1-i2-Myc-Ras bait fusion as well as the Myc-Ras fusion encoded by the empty bait expression vector (Figure 3.5). Although methionine can not completely suppress the expression of all the baits, it

suppresses the expression to a high level, which is not ideal but acceptable for screening to eliminate false positives. The potential problem might be that if there is an interaction between the bait and prey, there might be a low level of growth of the yeast cells on the Gal-L-U+4M plate where there should not be any growth at all.

The temperature sensitivity of the yeast strain is crucial in the RRS screening experiments. The yeast cells will not grow on the selective media at the restrictive temperature if there is no interaction between the bait and prey. However, if the yeast cells lost the temperature sensitivity through mutation, they will grow on the selective media at the restrictive temperature even if there is no interaction present. This will generate a large number of false positives and reduce the sensitivity of the experiment. The yeast cells must be tested for temperature sensitivity both before and after transformation, prior to any assays for protein-protein interactions. In this project, temperature sensitivity test confirmed that the bait-containing yeast cells were temperature sensitive (Figure 3.6).

The baits were also tested for autoactivation of the RRS. The i1-i2 bait did not autoactivate the RRS, but the i3 and Cter baits did autoactivate the RRS at the restrictive temperature (Figure 3.7). One possibility was that i3 and Cter contain potential membrane binding regions which lead to their binding to the membrane in the absence of preys. The binding of the bait to the membrane enabled the mRas to complement the mutated yRas, which led to yeast growth independent of the bait-prey interaction at the restrictive temperature. Alternatively these baits may somehow activate the Ras signalling pathway downstream of mRas and thus give rise to cell growth. There was not sufficient time to investigate this further however future work could be done by dividing i3 and Cter into several shorter regions and using each region as individual bait. The baits which would not autoactivate the RRS could be used in the cDNA library screening for interactors

Chapter 4 RRS screening and the identification of potential GCR1 interactors

4.1 Introduction

4.1.1 Library screening for GCR1 interactors using the RRS

A variety of methods are available for the identification of protein-protein interactions as describe in section 1.2. The RRS was chosen for use in this project as it offers several advantages over other methods. As with all yeast two-hybrid systems, it is a relatively cheap system with a high degree of sensitivity. It is large scale and high-throughput, so prey expression libraries can be screened for potential protein interactions with relative ease and speed. It can also be used to analyse known protein interactions and to map the domain that is responsible for the interaction. In addition, the process is done *in vivo* which is a natural environment for proteins and will theoretically achieve more accurate results than *in vitro* methods. Another appealing feature is that the identification of an interacting protein implies that at the same time the corresponding gene is cloned (Criekinge and Beyaert, 1999). Since the *Arabidopsis* genome has been fully sequenced, the identity of the gene can be easily found out by sequencing and blast searching again the *Arabidopsis* genome database.

Apart from the above advantages, there is one major advantage of the RRS compared to the traditional GAL4 system for protein-protein interaction analysis. The GAL4 system requires the transcriptional regulation that is induced in the nucleus of yeast by protein–protein interactions, it is not inherently suited to analyse membrane proteins, such as GCR1, which are unable to fold correctly in the nucleus (Milligan and White, 2001). However, in the RRS the interaction occurs at the cytoplasmic face of the membrane rather than in the nucleus, providing a more favorable environment for the

assembly of membrane proteins. The RRS has been used extensively to identify novel interaction partners, also in cases where common systems failed (Kruse *et al*., 2006). It is a valuable alternative interaction screening method that can widely be used among protein classes as demonstrated by several research groups (Broder *et al*., 1998; Kohler and Muller, 2003; Kruse *et al*., 2006). Taken the above points into account, it was deemed worthwhile to use the RRS to screen an *Arabidopsis* cDNA library for GCR1 interactors.

4.1.2 Research objectives and experimental approach

The objectives of this part of the project were:

(1) Screen an *Araibdopsis* cDNA library using the i1-i2 bait for GCR1 interactors

(2) Investigate the identity of the potential interactors using a PCR-Sequencing-Dot-blot based approach as illustrated in figure 4.1.

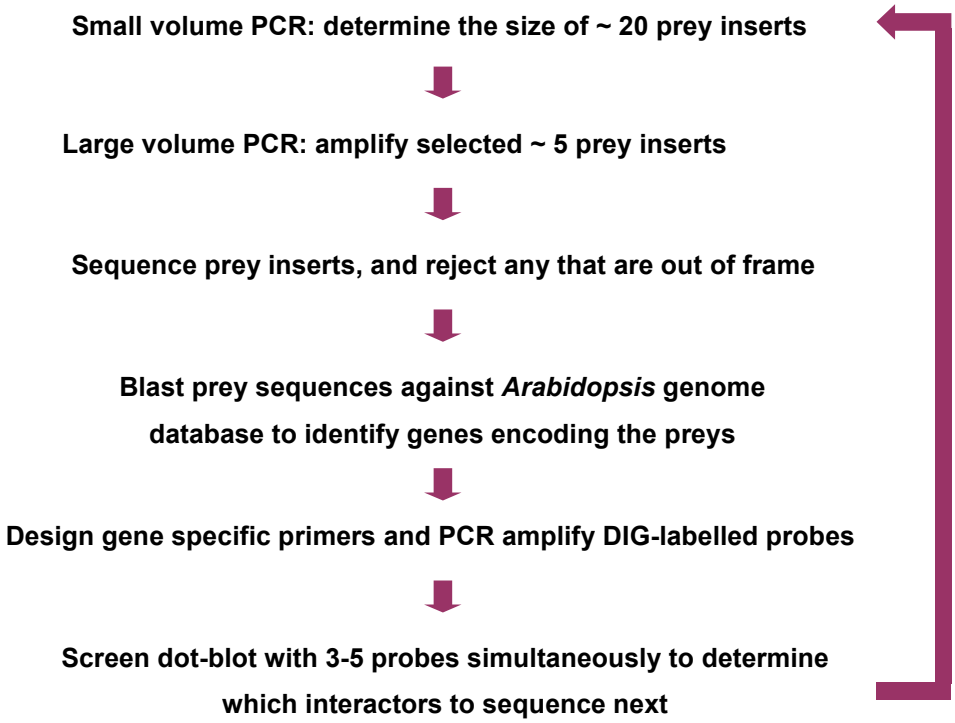


Figure 4.1 Flow chart of a PCR-Sequencing-Dot-blot based approach for the identification of potential interactors.

4.2 Results

4.2.1 Screening an *Arabidopsis* cDNA library for GCR1 interactors

About $1.3x10^6$ independent transformants of an *Arabidopsis* cDNA library were screened using the i1-i2 bait for GCR1 interactors as described in section 2.2.8. The yeast colonies were allowed to grow at 24°C until they had grown sufficiently to be replica plated. They were then replica plated onto the selective medium Gal-L-U-M, which allows the expression of both bait and prey proteins, at restrictive temperature 36°C. The interaction of the bait with the prey brings mRas to the cell membrane where it can interact with other signalling molecules that lead to cell proliferation and growth. Therefore colonies that appear on the selective medium at 36°C might contain potential interactors. About 1200 clones grew on the selective medium at 36°C.

To ensure that autoactivation of mRas or reactivation of the yRas, for example through a temperature sensitive reversion mutant, was not the cause of any growth seen, the yeast were subjected to a round of more stringent screening. The 1200 yeast colonies were re-streaked onto Glu-L-U+M medium, incubated at 24°C until sufficient growth had been obtained, and then replica plated onto the Gal-L-U-M medium and three control media, Glu-L-U-M, Gal-L-U+4M and YPD, which represses the expression of prey, bait, or both bait and prey respectively. After incubation at 36°C for 5-7 days, 774 clones (No.1 – No. 774) grew on the selective medium but not on the control media. These clones might be GCR1 interactors, so they were referred henceforth as potential interactors. The clones that grew on both selective medium and any of the control media were deemed to be false positives. Examples of interactions seen in the library screening using the i1-i2 bait are shown in figure 4.2.

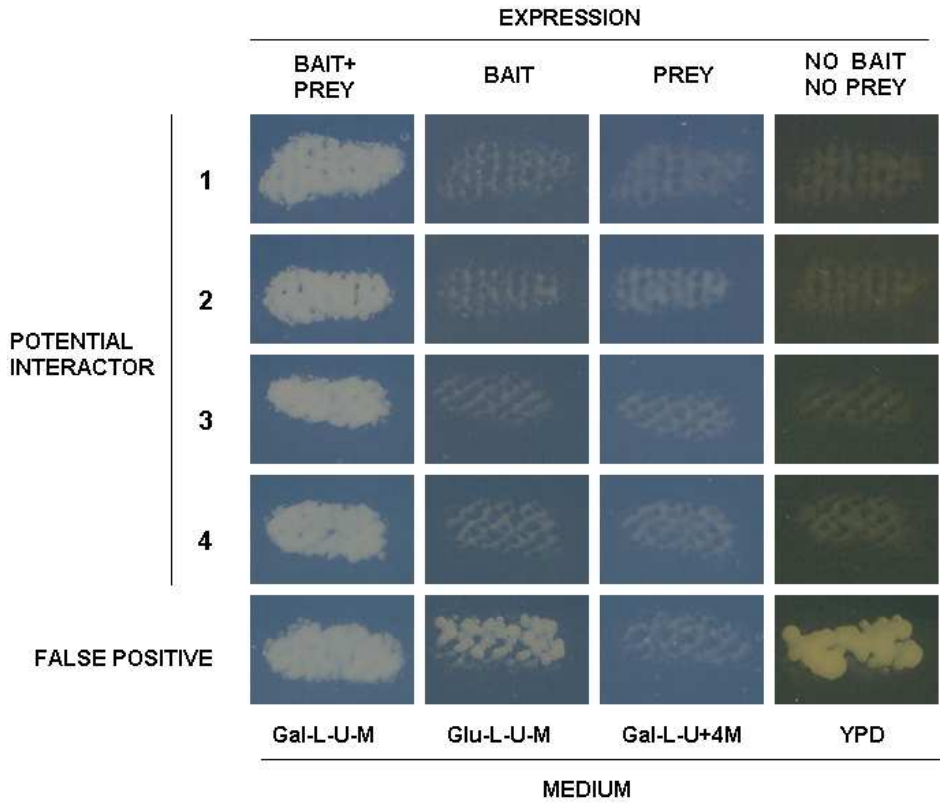


Figure 4.2 Examples of interactions seen in the RRS screen using the i1-i2 bait.

4.2.2 Identification of potential interactors using a PCR-Sequencing-Dot-blot based approach

4.2.2.1 First round identification by PCR-Sequencing-Dot-blot

Yeast colony PCR was performed for the first 58 clones (No.1-58), using vector specific T7 and pUraR as forward and reverse primer respectively (section 2.2.9, appendix 3, figure 4.3). According to the sizes of the PCR products, these clones were put into 11 groups each with near identical sizes. On the assumption that products with near identical sizes were probably identical clones, the PCR products of one representative for each group were purified and sequenced. The sequences were BLASTN searched against the *Arabidopsis* genome database to identify genes encoding the potential interactors. These initial 11 potential interactors are: TRX3 (thioredoxin H-type 3, At5g42980), TRX4 (thioredoxin H-type 4, At1g19730), PRMT (Protein

arginine N-methyltransferase, At4g29510), CL12A (50S ribosomal protein L12-1, chloroplast, At3g27830), CL12C (50S ribosomal protein L12-3, chloroplast, At3g27850), PRL27aC (60S ribosomal protein L27A, At1g70600), PAG1 (20S proteasome alpha subunit G, At2g27020), MLP168 (MLP-like protein 168, BetVI allergen family protein, At1g35310), AGL42 (MADS-box protein, At5g62165), ASK1 (E3 ubiquitin ligase SCF complex subunit SKP1/ASK1, At1g75950), AHB2 (non-symbiotic hemoglobin 2, At3g10520) (Table 4.1).

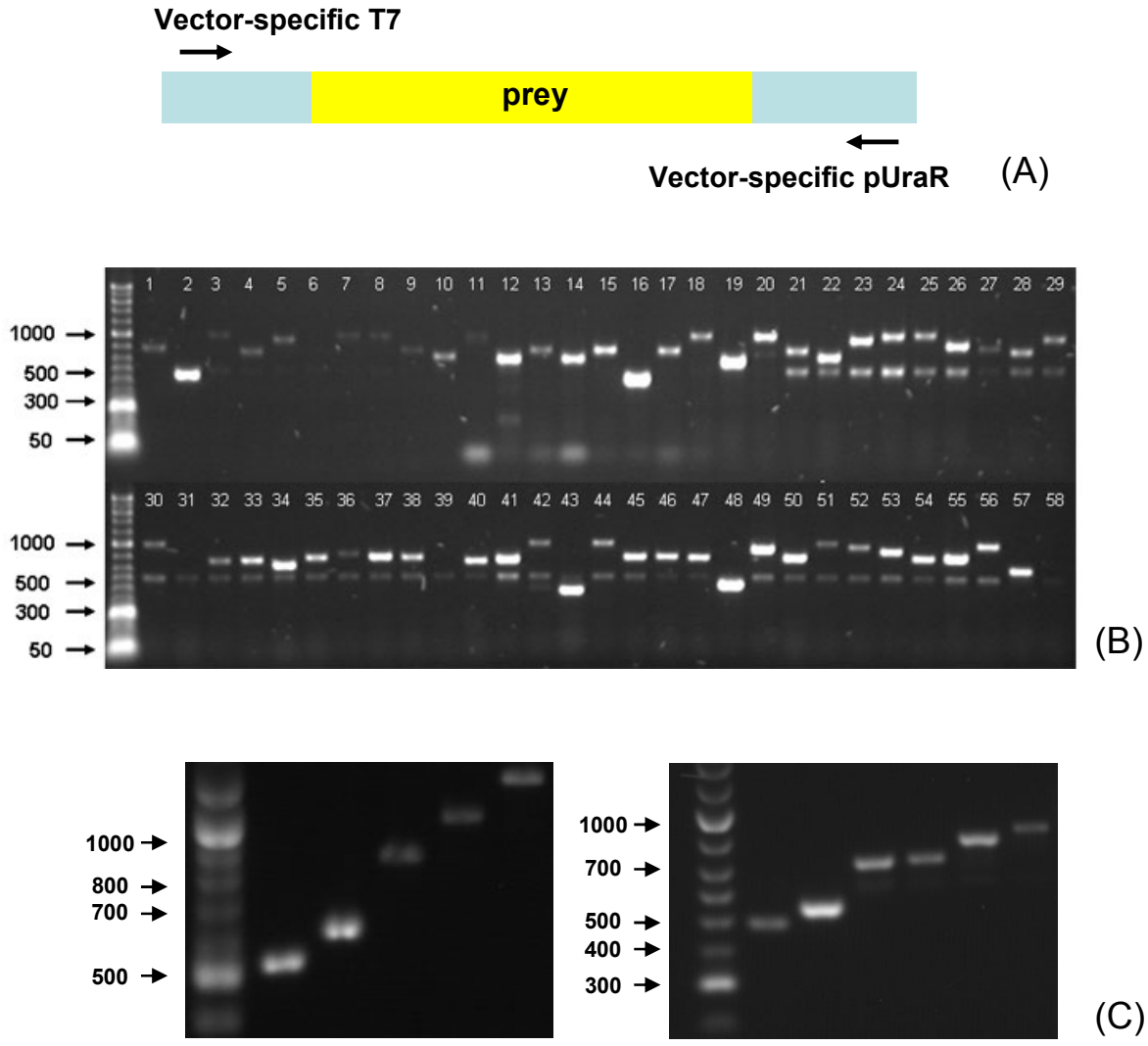


Figure 4.3 Example of yeast colony PCR. (A) PCR was performed using the cDNA of individual interactor as template and using T7 as forward and pUraR as reverse primer.

(B) Yeast colony PCR was performed for the first 58 clones. (C) Large volume PCR (50µl reaction) was carried out to amplify the representatives for purification and sequencing.

To work out the identities of the un-sequenced clones among the first 58 clones, 11 rounds of PCR (figure not shown) were carried out on the 58 clones using a gene specific forward primer, for instance TRX4Beg, and the pUraR reverse primer (appendix 3). The number of clones that produced PCR products for each of the 11 potential interactors is listed in table 4.1. The representing frequency of each potential interactor is also calculated, since the higher the represent frequency is, the more like that it is a true interactor.

Table 4.1 Potential interactors identified from clone No.1 to No.58.

Interactors	Numbers	Frequency
TRX3 (thioredoxin H-type 3)	14	36.2%
TRX4 (thioredoxin H-type 4)	7	
PRMT (Protein arginine N-methyltransferase)	1	1.7%
CL12A (50S ribosomal protein L12-1, chloroplast)	20	48.3%
CL12C (50S ribosomal protein L12-3, chloroplast)	7	
PRL27aC (60S ribosomal protein L27A)	1	
ASK1 (E3 ubiquitin ligase SCF complex subunit SKP1/ASK1)	3	5.2%
AHB2 (non-symbiotic hemoglobin 2)	1	1.7%
PAG1 (20S proteasome alpha subunit G)	2	3.4%
AGL42 (MADS-box protein)	1	1.7%
MLP168 (MLP-like protein 168, BetVI allergen family protein)	1	1.7%

The 11 potential interactors identified from clones No.1-58 are listed in the first column. The numbers of clones that produced PCR products for each of the 11 potential interactors is listed in the middle column. The frequency in the last column shows how frequently each interactor turned up.

The identities of the rest of the 774 clones could be worked out by colony PCR and sequencing in the same way as described above. However, it would require a large amount of time, materials and money to complete the work. Since there were clones that are represented more than once in the first 58 clones, there would be clones that are represented more than once in the rest

of the 716 unidentified clones. Therefore, dot-blots were carried out to identify the clones appeared more than once. Because the thioredoxins and the ribosomal proteins were the majority (84.5%) of the first 58 identified putative interactors, dot-blots using them as probes would pick up a large number of the same clones that exist in the 716 unidentified clones (No.59-774).

The TRX3, TRX4, CL12C and PRL27aC probes were DIG-labeled by PCR, before being purified from 1.2% agarose gel (section 2.2.11.2, figure 4.4,). The predicted sizes of the PCR products are 357 bp, 360 bp, 301 bp, and 313 bp respectively. The actual sizes of the PCR products obtained were larger than the sizes predicted, however, this is because the dUTPs were DIG-labelled and the DNA length markers were not. Eight nylon membranes which contain the cDNAs of the unidentified 716 putative interactors were prepared as described in section 2.2.11.1. Each of the No.1 to No.7 membranes contains 96 samples, and membrane No.8 contains 44 samples.

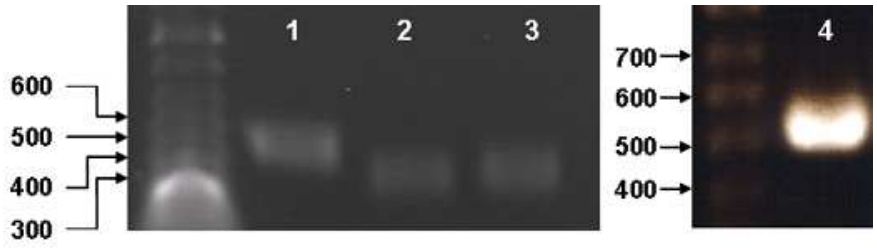


Figure 4.4. PCR synthesis of DIG-labelled probes. lane 1: TRX3, 357 bp; lane 2: CL12C, 301 bp, lane 3: PRL27aC, 313 bp, lane 4: TRX4, 360 bp. The actual size of each PCR product was larger than the size compared with DNA length markers, because the dUTPs were DIG-labelled.

If the probes were used individually to figure out how frequently individual clones turned up, four dot-blots would have to be carried out for each membrane (32 dot-blots in total). It would take too much time to finish the work. Instead, we used pooled probes for the dot-blots so that only one dot-blot was conducted for each membrane (8 dot-blots in total). After the dot-blot screening (Figure 4.5), 431 (60.2%) clones were identified using these pooled probes and 285 (39.8%) cloned remained unidentified.

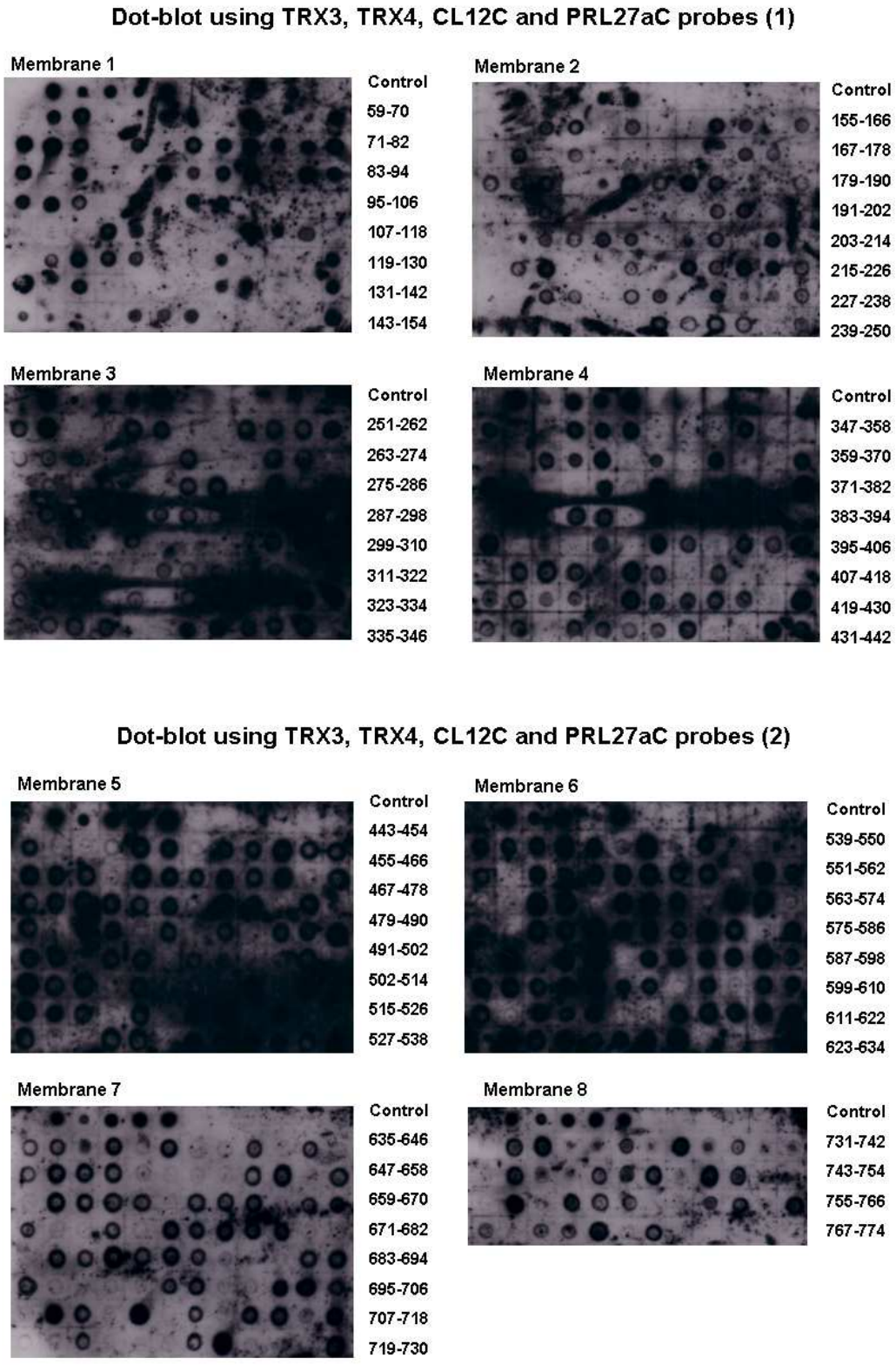


Figure 4.5 First round dot-blot identification of clones represented more than once. Dot-blots were performed for the 716 unidentified clones (No. 59-774) using the pooled probes: TRX3, TRX4, CL12C and PRL27aC. The controls (dot 1 to 6 in row one) are the cDNA of GCR1, TRX3, TRX4, CL12A, CL12C and PRL27aC respectively.

4.2.2.2 Second round identification by PCR-Sequencing-Dot-blot

To identify more clones, yeast colony PCR was performed for some more unidentified clones, using T7 as forward and pUraR as reverse primer. According to the sizes of PCR products, the interactors were in six different groups. One representative from each group was sequenced. These six potential interactors are: MLP-like protein 43 (MLP43, At1g70890), putative plant defensin-fusion protein (PDF2.2, At2g02100), C2 domain-containing protein (C2dcp, At2g01540), glycine-rich protein (Glyrich, At5g46730), ADP-glucose pyrophosphorylase family protein (ADPGlc PPase, At1g74910) and cytochrome P450 (P450, At3g10570). The second round dot-blots were performed using these potential interactors as probes to identify the same interactors in the rest of the unidentified 285 colonies. The probes were made in the same way as described above (Figure 4.6). The eight membranes used in the first round of dot-blots were stripped to remove the previous probes and re-probed with these new probes (section 2.2.11.4)

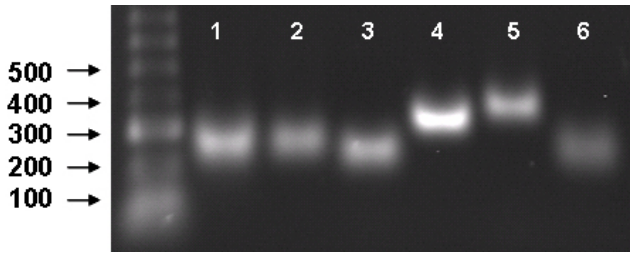
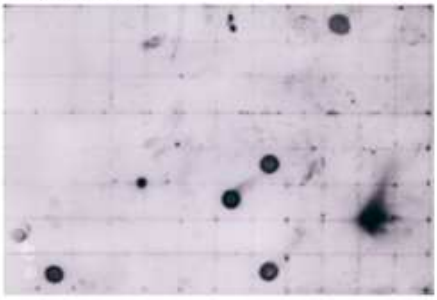


Figure 4.6 PCR synthesis of DIG-labelled probes. lane 1: Glyrich, 142 bp; lane 2: ADPGlc PPase, 147 bp; lane 3: P450, 126 bp; lane 4: MLP43, 206 bp; lane 5: PDF2.2, 240 bp; lane 6: C2dcp, 189 bp. The actual size of each PCR product was larger than the size compared with DNA length markers, because the dUTPs were DIG-labelled.

In this round of screening (Figure 4.7), 63 (22.1%) clones were identified as MLP43, PDF2.2, C2dcp, Glyrich, ADPGlc PPase and P450. However, there were still 222 (77.9%) not identified.

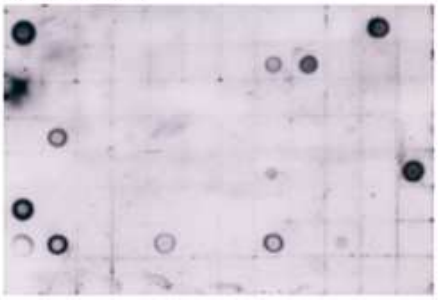
Dot-blot using MLP43, PDF2.2 and C2dcp probes (1)

Membrane 1



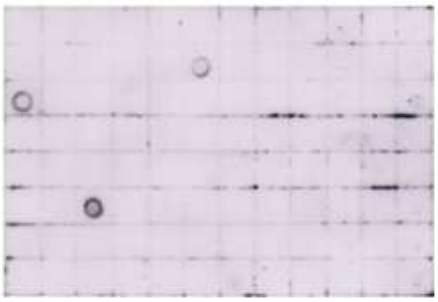
59-70
71-82
83-94
95-106
107-118
119-130
131-142
143-154

Membrane 2



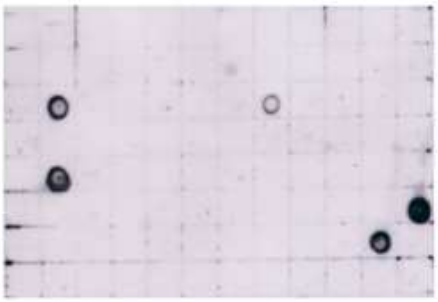
155-166
167-178
179-190
191-202
203-214
215-226
227-238
239-250

Membrane 3



251-262
263-274
275-286
287-298
299-310
311-322
323-334
335-346

Membrane 4



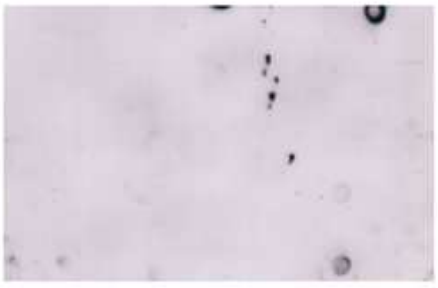
347-358
359-370
371-382
383-394
395-406
407-418
419-430
431-442

Dot-blot using MLP43, PDF2.2 and C2dcp probes (2)

Membrane 5

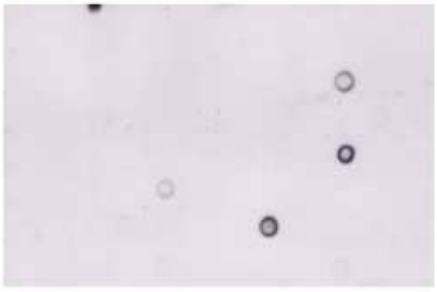
443-454
455-466
467-478
479-490
491-502
502-514
515-526
527-538

Membrane 6



539-550
551-562
563-574
575-586
587-598
599-610
611-622
623-634

Membrane 7



635-646
647-658
659-670
671-682
683-694
695-706
707-718
719-730

Membrane 8

731-742
743-754
755-766
767-774

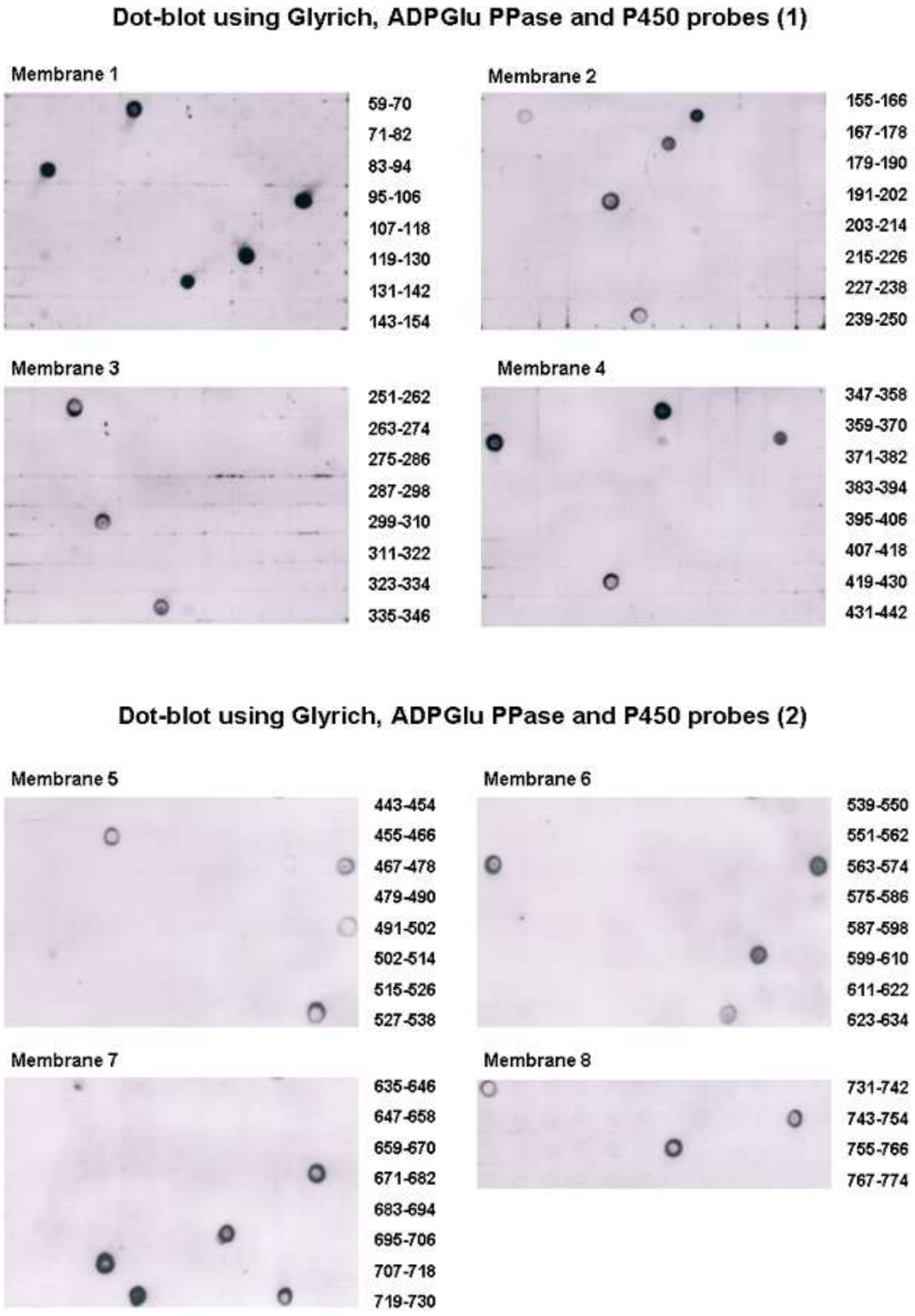
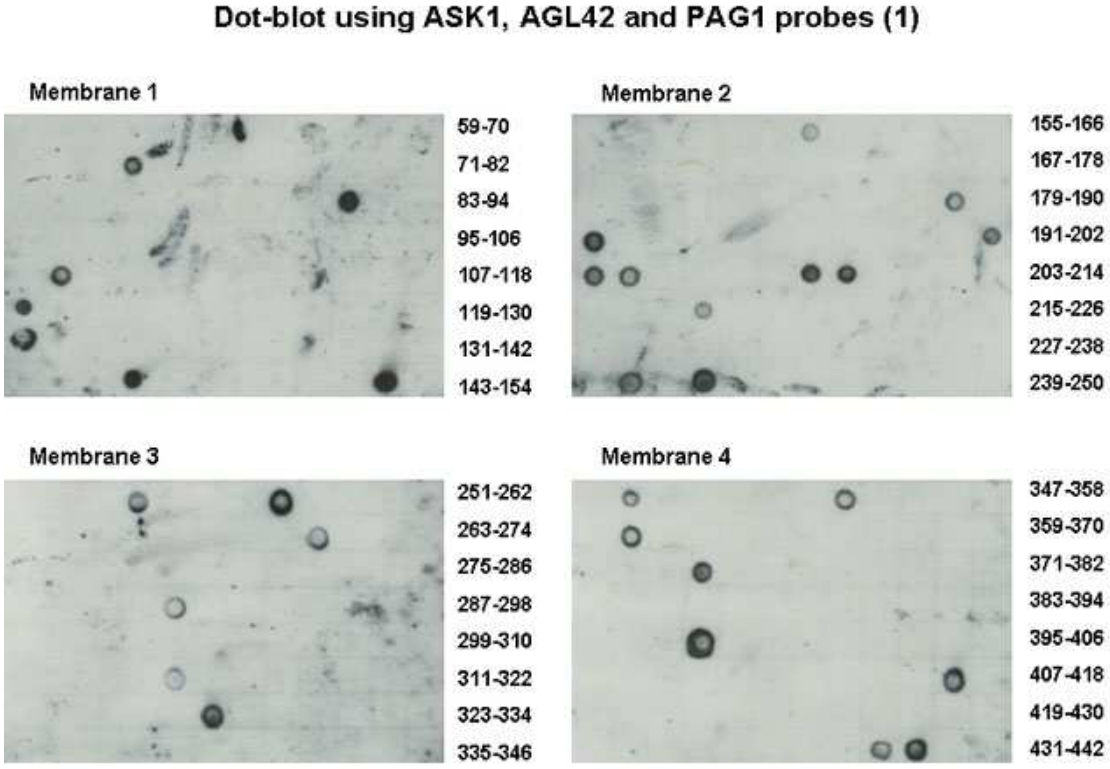


Figure 4.7 Second round dot-blot identification of clones represented more than once. Dot-blots were performed for the 716 clones (No. 59 – 774). Two groups of pooled probes were used, one group contains the MLPR, PDF2.2, C2dcp probes, and the other group contains the Glyrich, ADPglc PPase and P450 probes.

4.2.2.3 Further rounds of identification by PCR-Sequencing-Dot-blot

A few more rounds of PCR-Sequencing-Dot-blot screening were carried out (Figure 4.8-4.12). After the third round of dot-blots (Figure 4.8), only 169 (23.6%) clones were left unidentified. The cDNAs of these clones were reloaded to three new small membranes (Table 4.2) to reduce the total membrane areas to be screened, which led to reducing the screening time by more than 75%.



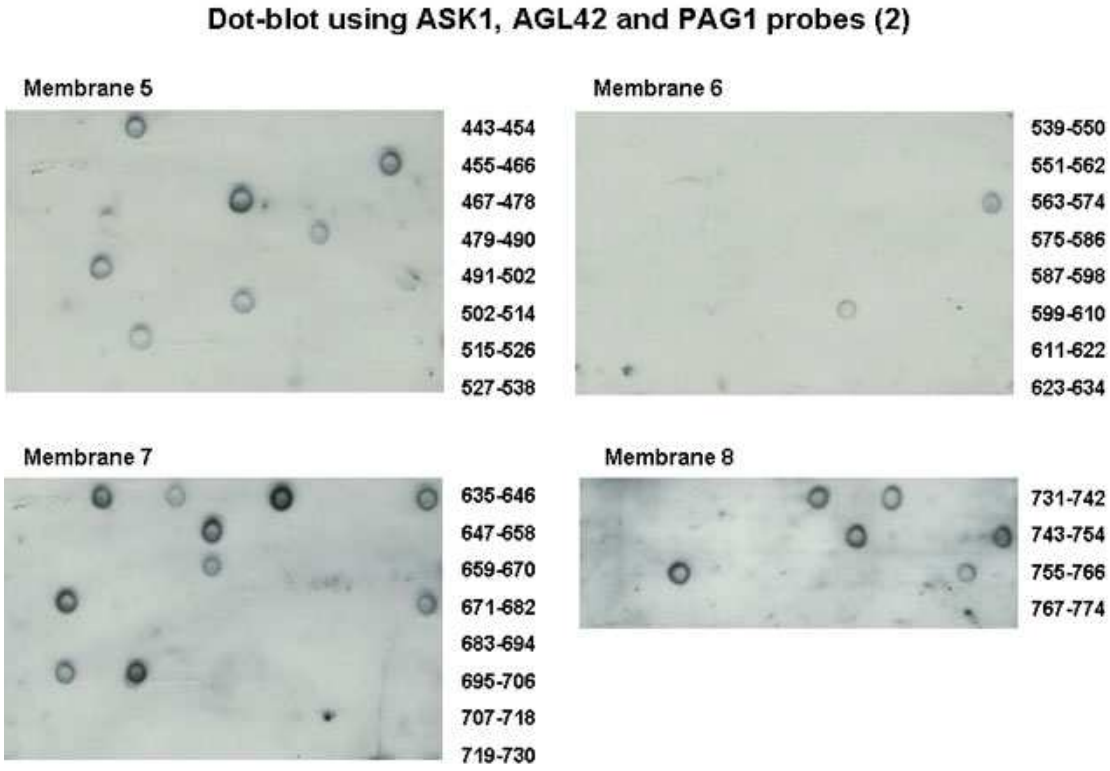


Figure 4.8 Third round dot-blot identification of clones represented more than once. Dot-blots were performed for the 716 clones (No. 59 – 774). Pooled probes containing ASK1, AGL42 and PAG1 were used.

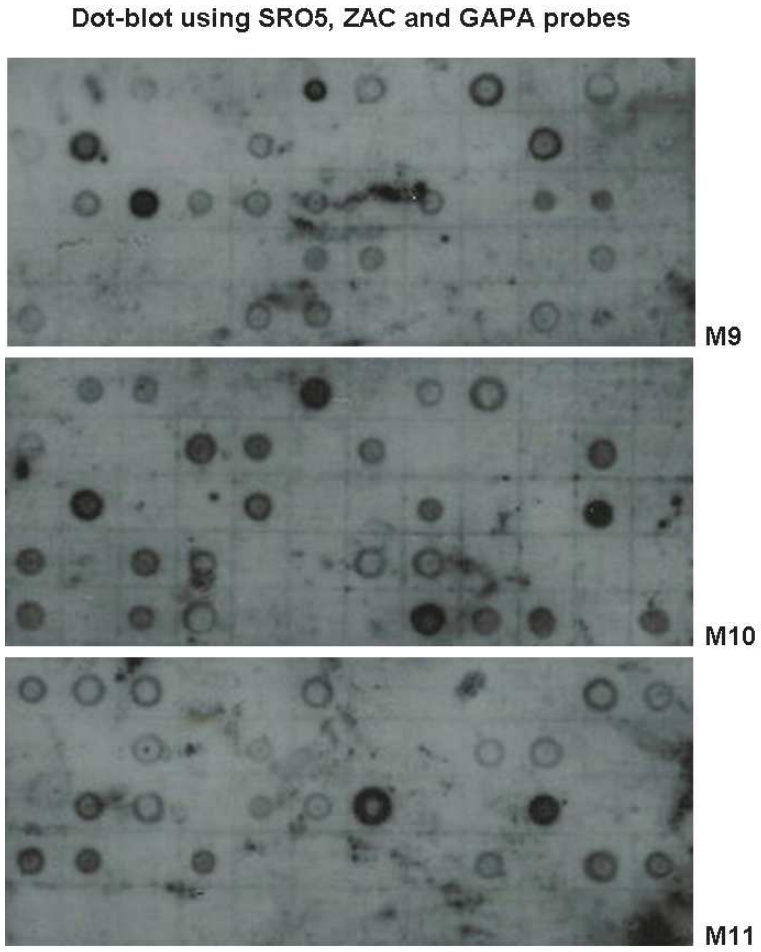


Figure 4.9 Forth round dot-blot identification of clones represented more than once. Dot-blot was performed using pooled probes containing SRO5, ZAC and GAPA for membrane (M) 9 -11.

Table 4.2 Yeast colonies loaded on Membrane 9 – 11

Membrane		Clone No.
M9	Row1	59, 62, 65, 66, 69, 76, 86, 87, 98, 100, 103, 104
	Row2	106, 107, 108, 109, 112, 113, 118, 124, 128, 129, 132, 134
	Row3	135, 136, 140, 141, 145, 151, 152, 159, 167, 169, 171, 172
	Row4	178, 182, 188, 194, 196, 197, 198, 201, 212, 219, 221, 239
	Row5	241, 243, 249, 257, 266, 270, 274, 277, 279, 280, 283, 285
M10	Row1	287, 295, 296, 298, 299, 304, 305, 306, 307, 312, 319, 321
	Row2	326, 331, 335, 339, 346, 349, 355, 357, 358, 366, 368, 373
	Row3	376, 379, 381, 383, 388, 394, 397, 400, 403, 407, 414, 416
	Row4	433, 444, 445, 449, 480, 483, 485, 496, 499, 500, 506, 508
	Row5	510, 512, 514, 520, 523, 525, 526, 528, 529, 532, 544, 546
M11	Row1	548, 550, 552, 560, 564, 567, 571, 576, 585, 587, 589, 592
	Row2	612, 616, 617, 623, 633, 641, 644, 651, 653, 654, 657, 659
	Row3	673, 675, 683, 690, 691, 697, 703, 707, 712, 717, 719, 720
	Row4	722, 724 727, 729, 734, 735, 741, 742, 743, 746, 755, 762
	Row5	772

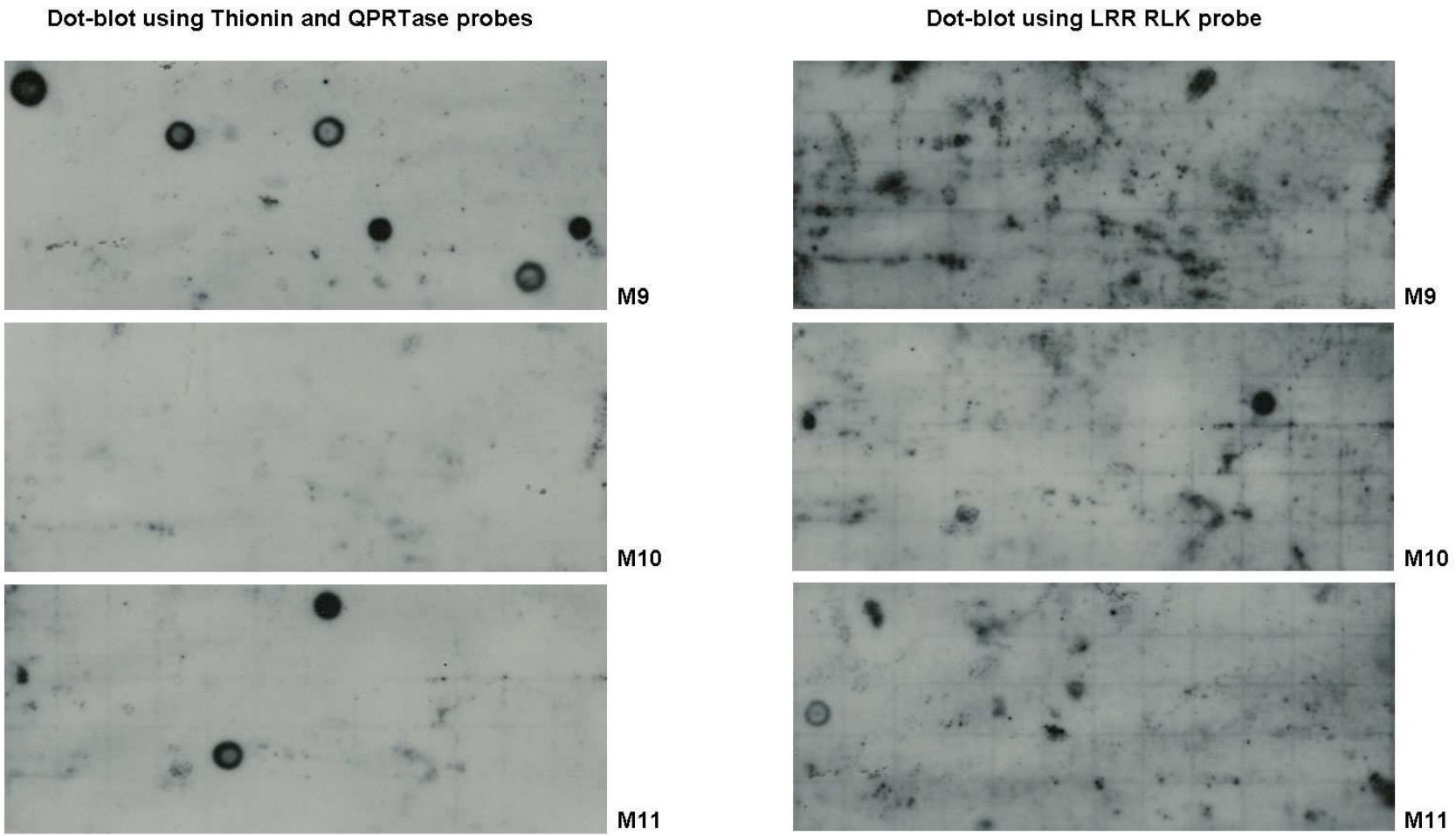


Figure 4.10 Fifth round dot-blot identification of clones represented more than once. Dot-blots were performed using two groups of pooled probes for membrane 9 -11. One group contains the Thionin and QPRTase probes, and the other group contains the LRRRLK probe.

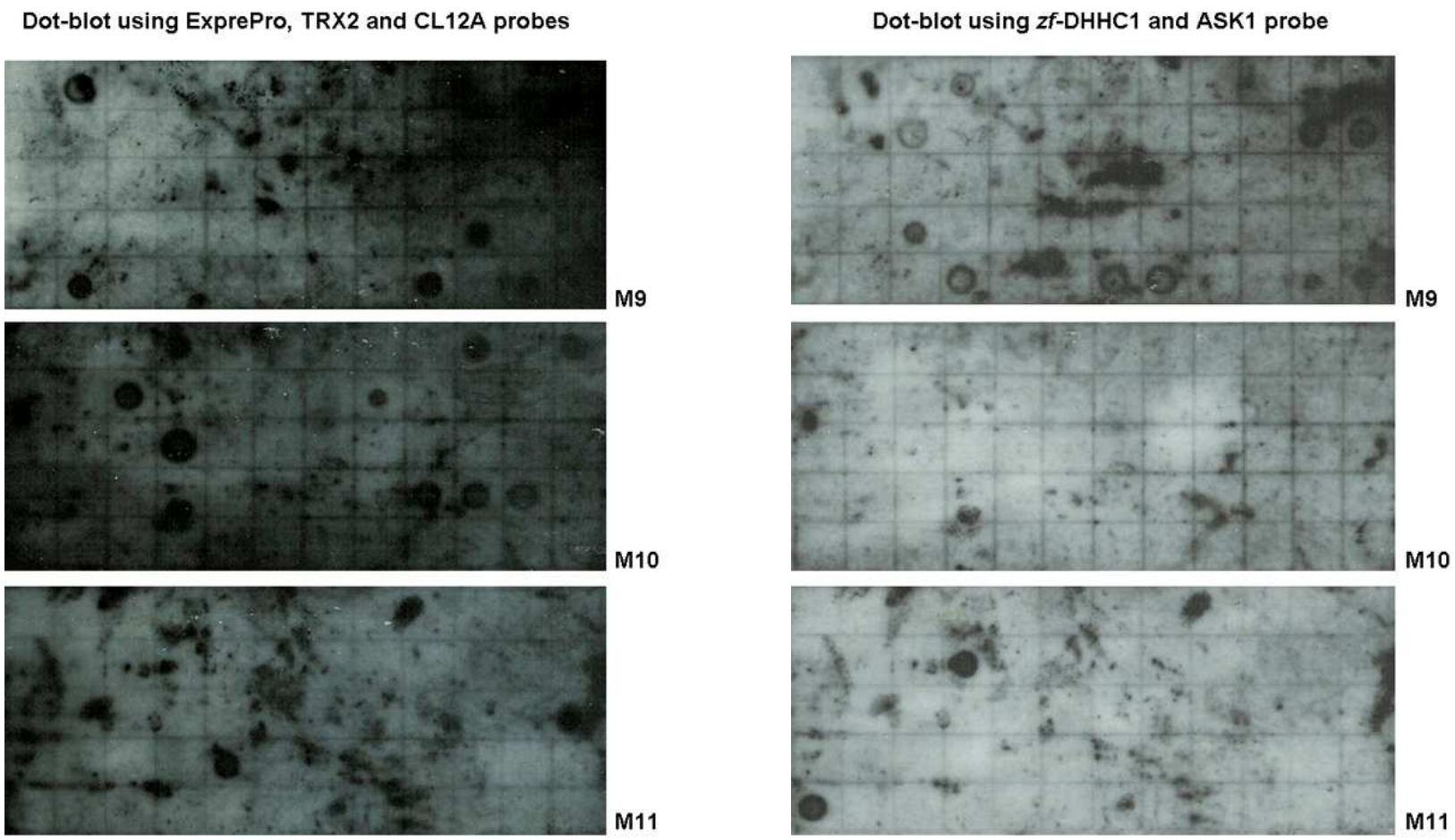


Figure 4.11 Sixth round dot-blot identification of clones represented more than once. Dot-blots were performed using two groups of pooled probes for membrane 9 -11. One group contains the ExprePro, TRX2 and CL12A probes, and the other group contains the *zf*-DHHC1 and ASK1 probes.

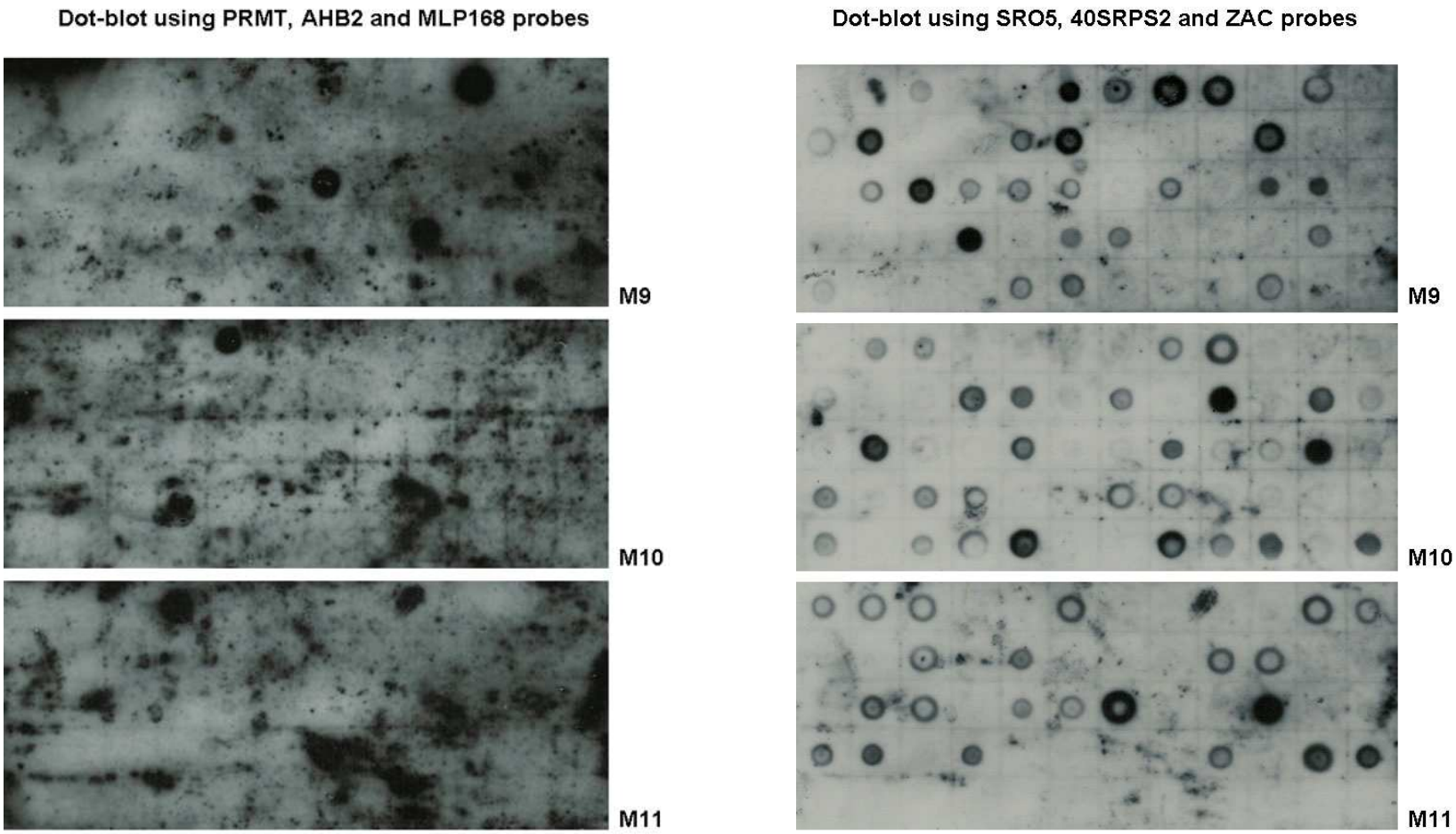


Figure 4.12 Seventh round dot-blot identification of clones represented more than once. Dot-blots were performed using two groups of pooled probes for membrane 9 -11. One group contains the PRMT, AHB2 and MLP168 probes, and the other group contains the SRO5, 40SROS2 and ZAC probes.

In total, 7 rounds of dot-blots were carried out and 27 genes encoding preys were found to be represented in 728 interactors out of the total of 774 interactors. Table 4.3 is a list of these 27 genes. By this stage, only 46 clones have not been identified. Considering that each round of PCR-Sequencing-Dot-blot took about 3-4 weeks to complete and the schedule was tight, we decided to leave investigation of the 46 remaining unidentified clones to the future and move on to the next stage of the project.

Table 4.3 Identified potential interactors

	AGI No.	Abbre.	Full Name	Identity	In frame	Full length
1	At5g39950	TRX2	Thioredoxin H-type 2	100%	Yes	No
2	At5g42980	TRX3	Thioredoxin H-type 3	100%	Yes	Yes
3	At1g19730	TRX4	Thioredoxin H-type 4	99%	Yes	Yes
4	At3g10520	AHB2	Non-symbiotic hemoglobin 2	100%	Yes	No
5	At3g51390	*zf*-DHHC1	Zinc finger (DHHC type) family protein	100%	Yes	Yes
6	At2g01540	C2dcp	C2 domain-containing protein	96%	Yes	No
7	At4g21160	ZAC	Zinc finger and C2 domain protein	100%	Yes	No
8	At2g02780	LRR RLK	Leucine-rich repeat transmembrane protein	99%	Yes	No
9	At1g75950	ASK1	E3 ubiquitin ligase SCF complex subunit, SKP1	99%	Yes	Yes
10	At2g27020	PAG1	20S proteasome alpha subunit G	100%	Yes	No
11	At4g29510	PRMT	Protein arginine N-methyltransferase	100%	Yes	No
12	At3g10570	P450	Cytochrome P450	76%	Yes	No
13	At1g74910	ADPGlc PPase	ADP-glucose pyrophosphorylase family protein	99%	Yes	No
14	At2g01350	QPRTase	Quinolinate phosphoribosyl transferase family protein	100%	Yes	No
15	At5g62165	AGL42	AGAMOUS LIKE 42 (MADS-box protein)	98%	Yes	No

16	At1g35310	MLP168	MLP-like protein 168 (BetVI allergen family protein)	96%	Yes	Yes
17	At1g70890	MLP43	MLP-like protein 43 (Major latex protein-related)	100%	Yes	No
18	At3g08030	ExprePro	Expressed protein, contains Pfam profile PF04862	100%	Yes	No
19	At3g27830	CL12A (RPL12-A)	50S ribosomal protein L12-1, chloroplast	99%	Yes	Yes
20	At3g27850	CL12C (RPL12-C)	50S ribosomal protein L12-3, chloroplast	97%	Yes	No
21	At1g70600	RPL27aC	60S ribosomal protein L27A	99%	No	No
22	At1g58983	40SRPS2	40S ribosomal protein S2 (At1g58684, At1g59359)	99%	No	No
23	At2g02100	PDF2.2	Putative plant defensin-fusion protein	72%	No	Yes
24	At1g72260	Thionin	Thionin	100%	No	Yes
25	At5g46730	Glyrich	Glycine-rich protein	99%	No	No
26	At5g62520	SRO5	Encodes a protein with similarity to RCD1 but without the WWE domain. Its presence suggests a role for the protein in ADP ribosylation	99%	No	No
27	At3g26650	GAPA	Glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	94%	No	No

4.2.3 Bioinformatics analysis of the potential interactors

Of the 27 potential interactors, 7 are not in frame to the myristoylation sequence which is located upstream of the cDNA sequence (Figure 2.1) and were therefore rejected. Bioinformatics analysis was carried out on the 20 remaining potential interactors to investigate possible biological relevance to GCR1. The function of each potential interactor is summarised in table 4.5.

Table 4.5 Summary of the functions of the 20 potential interactors

Thioredoxins (TRX2, TRX3 and TRX4) (1) Thioredoxins reduce disulfide bridges of target proteins by the reversible formation of a disulfide bridge between two neighboring cysteine residues present in the conserved site WCXPC (Kumar *et al*., 2004). (2) They are involved in oxidative stress response, e.g. the detoxification of H_2O_2 via the enzyme thioredoxin peroxidase or peroxiredoxin (Andoh *et al*. 2002). (3) In plants, they act as a reducing system participating in the mobilization of protein reserves during seed germination (Alkhalfioui *et al*., 2007). They are potent regulators of membrane-bound, receptor-like kinases and have been shown to interact with the C terminus of the S gene product and are involved in self-incompatibility (Bower *et al*., 1996; Schurmann and Jacquot (4) They modulate the DNA binding activity of some transcription factors (Spyrou *et al*., 1997) and also play a role as chaperones in protein folding (Kern *et al*., 2003; Jurado *et al*., 2006).
AHB2 (1) Hemoglobins are ubiquitous proteins in most organisms (Arredondo-Peter *et al*., 1998). (2) Plants have two types of hemoglobins, symbiotic-type and nonsymbiotic-type. Symbiotic hemoglobins facilitate oxygen diffusion from outside the root to nitrogen-fixing endosymbiotic bacteria of legumes (Kundu *et al*., 2003). The physiological role of non-symbiotic hemoglobins is unknown. (3) They potentially could be involved in stress responses by sequestering oxygen or nitrogen oxide (Perazzolli *et al*., 2004)
***zf*-DHHC1** (1) Zinc finger proteins in plants have been identified to be involved in transcriptional regulation, floral organogenesis, leaf initiation, lateral shoot initiation, gametogenesis and stress response (Englbrecht *et al*., 2004). (2) The DHHC zinc finger domain has been predicted to be involved in protein-protein or protein-DNA interactions (TAIR, 2005). (3) The DHHC protein family palmitoylate protein targets (Roth *et al*., 2002; Resh 2006). Protein palmitoylation regulates membrane tethering for key proteins in cell signalling and membrane trafficking. GPCRs and other TM proteins are palmitoylated on one or several cysteine residues near the TM domains (Qanbar and Bouvier, 2003; Smotrys and Linder, 2004).
C2dcp C2 domain is a Ca^{2+}-dependent membrane-targeting module that is typically found in multi-domain proteins involved in signal transduction (*e.g*., PKC, cytosolic PLA_2 ($cPLA_2$), phospholipases C (PLC), plant phospholipase D (PLD), and phosphatidylinositol 3-kinase) or membrane trafficking (*e.g*., synaptotagmins, rabphilin-3A, and Unc-13) (Rizo and Sudhof, 1998; Cho, 2001; Jensen *et al*., 2000).
ZAC (1) It is associated both with the plasma membrane and the Golgi apparatus. It is expressed to the highest levels in flowering tissue, rosettes and roots (Jensen *et al*., 2000). (2) ZAC-zinc finger domain shows significant similarity to ARF GAP proteins from animals and fungi (Jensen *et al*., 2000). (3) A recombinant ZAC possess GTPase-activating activity on *Arabidopsis* ARF proteins. A region of this ZAC binds high specifically to phosphatidylinositol-3-phosphate, which plays a crucial role in several signalling and membrane trafficking pathways (Jensen *et al*., 2000).
LRR RLK (1) The LRR RLKs contain an extracellular LRR and a Ser/Thr kinase domain, and are localized at the plasma membrane. In Arabidopsis, more than 200 LRR RLKs have been

identified and classified into 13 subfamilies, but most of their functions are unknown (Shiu and Bleecker, 2001). (2) The well characterized ones are *CLAVATA1 (CLV1)* which is involved in meristem differentiation, *BRASSINOLIDE INSENSITIVE 1 (BRI1)* and *BRI1 ASSOCIATED RECEPTOR KINASE 1 (BRK1)* which are involved in brassinosteroid perception, and FLAGELLIN SENSITIVE 2 (FLS2) which is involved in defense/pathogen recognition (Dievart and Clark, 2004).

ASK1

(1) ASK1 is a fundamental component of the SCF complex, which regulates a variety of processes in plants, such as flower development, circadian, gibberellin signaling, light signalling, defense response and leaf senescence (Zhao *et al*., 2003). (2) *ASK1* gene is essential for male meiosis. ASK1 and ASK2 play vital role in *Arabidopsis* embryogenesis and postembryonic development (Liu *et al*., 2004).

PAG1

PAG1 is the 20S proteasome alpha subunit G. The 20S proteasome has been reported to play a role in the secondary antioxidative defence (Reinheckel *et al*, 1998).

PRMT

(1) Protein arginine methylation is a common post-translational modification that has been implicated in signal transduction, RNA processing, transcriptional regulation, DNA repair and protein-protein interaction (Lee *et al*., 2005; Boisvert *et al*., 2003). (2) Type I protein arginine *N*-methyltransferases catalyze the formation of asymmetric *N*G,*N*G-dimethylarginine residues in proteins by transferring methyl groups from *S*adenosylmethionine (AdoMet) to the guanidino nitrogen atoms of arginine residues (Tang *et al*., 1998).

P450

(1) In plants, P450s are known to play important roles in production of hormones, pigments, oils, defensive compounds and endogenous lipophilic compounds (Mizutani *et al*, 1998; Nguyen *et al*., 2001). (2) Oxidative detoxification of a number of herbicides in plant tissues is also achieved by a P450 dependent monooxygenase system (Mizutani *et al*, 1998). (3) Cytochrome P450 enzymes of the closely related CYP90 and CYP85 families catalyze essential oxidative reactions in the biosynthesis of brassinosteroid (BR) hormones (Bancos *et al*, 2002).

ADPGlc PPase

(1) ADP-glucose is known to play a prominent role in starch metabolism in plants. ADP-glucose is synthesized by the enzyme ADP-glucose pyrophosphorylase (APGase) and is then polymerized via the soluble and granulebound starch synthases (McCoy *et al*., 2006). (2) Thioredoxin could reduce or oxidize the chloroplastic ADP-GlcPPase during the light/dark cycle, thus providing a fine-tuned regulation of starch synthesis on chloroplasts (Ballicora *et al*, 2000).

QPRTase

It contains Pfam profile: PF01729 quinolinate phosphoribosyl transferase, C-terminal domain. Quinolinate phosphoribosyl transferase (QPRTase) or nicotinate-nucleotide pyrophosphorylase is involved in the de novo synthesis of NAD in both prokaryotes and eukaryotes. It catalyses the reaction of quinolinic acid with 5-phosphoribosyl-1-pyrophosphate (PRPP) in the presence of Mg^{2+} to give rise to nicotinic acid mononucleotide (NaMN), pyrophosphate and carbon dioxide (TAIR, 2005).

AGL42

(1) AGL42 is a MADS-box protein. The MADS-box genes encode a family of transcription factors, which control developmental processes in flowering plants ranging from meristem

and organ identity, flowering time to root development (Aswath and Kim, 2005). (2) It has been shown that AGL42 is expressed in the quiescent center (QC) and maintains the structure and developmental function of the root meristem. Mutations in all enriched transcription factor genes including AGL42 exhibited no detectable root phenotype, raising the possibility of a high degree of functional redundancy in the QC (Nawy *et al*, 2005).
MLP168, MLP43, ExprePro Function unknown
Ribosomal proteins (CL12-A, CL12-C) Were assumed to be false positives because they are known to function in protein translation and are extremely abundant mRNAs

4.3 Discussion

4.3.1 Library screening

Ideally one would hope to screen for protein-protein interactions in plant-based systems. To date, it is difficult to investigate protein-protein interactions *in planta* on a large scale (Dortay *et al*., 2006) so for this reason we pursued an alternative approach. So far, the most widely used high-throughput screening of protein-protein interactions *in vivo* is the yeast two-hybrid system (Mukherjee *et al*., 2001). In this project, the RRS was used to screen an *Arabidopsis* cDNA library using the GCR1 derived i1-i2 bait. When screening a cDNA library, a good representation is crucial. In classical two-hybrid library preparations only one out of six fused cDNAs is in the correct frame, therefore, to screen an *Arabidopsis* cDNA library until saturation, more than $1.6x10^5$ (25,498 (genes) x 6 = 152,988 $\approx$ $1.6x10^5$) yeast transformants need to be screened. Taking this into account, $1.3x10^6$ transformants were screened using the i1-i2 bait. 20 potential interactors were represented in 728 transformants which showed reproducible growth on selective media at restrictive temperature. This indicated that the RRS is a sensitive and powerful system for studying novel protein-protein interactions at the inner face of the plasma membrane.

The interaction between GCR1 and the heterotrimeric G-protein subunits were not observed using the GAL4 system (Humphrey and Botella, 2001;

Pandey and Assmann, 2004). However, using a split-ubiquitin two-hybrid system, Pandey and Assmann (2004) demonstrated that GCR1 interacts with the G-protein α subunit GPA1. Therefore, an obvious protein partner for GCR1 that we might have expected to find in the library screening was GPA1. But this was not found as a GCR1 interactor in this screen. The most likely explanation is that the GPA1 interacts with i3 or Cter which did not function as suitable baits in the RRS, as an intact i3 and a free Cter of GCR1 have been shown to be required for the GCR1-GPA1 interaction (Pandey and Assmann, 2004). The other reason could be that the interaction is too transient to be detected by the RRS. As GCR1 is an integral membrane protein with multiple intracellular domains, the binding sites for its protein partners may contain many residues. It is best to carry out the screening with the entire protein but not the subfragments. This can be achieved by using rRRS with GCR1 as a membrane localised bait and performing screening with the reverse cDNA library.

Although it has been shown that the first 105 amino acids of GCR1 (from the beginning of GCR1 to the end of TM3) are not required for its interaction with GPA1 (Pandey and Assmann, 2004), accumulating evidence indicates that both i1 and i2 of mammalian GPCR contribute to the recognition and interaction with G-protein or its downstream effectors (Pin *et al*., 1994; Gomeza *et al*., 1996; Francesconi and Duvoisin, 1998; Wess, 1998; Yamashita *et al*., 2000; Liu and Wu, 2003; Geng *et al*., 2004; Rovati *et al*., 2007). Therefore, library screening using the i1-i2 bait which comprises the i1 and i2 regions of GCR1 are expected to identify a certain number of interactors. As shown in table 4.5 and table 4.6, twenty potential interactors with various functions were identified. Some of these can reasonably be expected to be false positives or artifacts of the system. For example, ribosomal proteins identified in the screen are almost certainly false positive and are present because mRNAs representing them are extremely abundant in most eukaryotic cells. Even though some candidates looked more promising than the others, it was hard to decide at this stage which ones

should be discarded and which ones to be worked on. As library screening by its nature is quantitative, the interaction identified in the initial screening should be subsequently re-examined by extracting the prey plasmid from the original yeast colony and re-transforming it back to the yeast cells with bait plasmid to test whether the interaction can still be observed. No further investigation of the nature of their interactions should be carried out prior to the re-examination (chapter 5).

Table 4.6 Summary of the 20 potential interactors according to function

Category	Interactor	AGI No.
Reactive oxyen species and stress responses	TRX2	At5g39950
	TRX3	At5g42980
	TRX4	At1g19730
	AHB2	At3g10520
Membrane targeting or signal transduction related	*zf*-DHHC1	At3g51390
	C2dcp	At2g01540
	ZAC	At4g21160
	LRR RLK	At2g02780
Ubiquitin proteasome related	ASK1	At1g75950
	PAG1	At2g27020
Enzymes	PRMT	At4g29510
	P450	At3g10570
	ADPGlc PPase	At1g74910
	QPRTase	At2g01350
Others	AGL42	At5g62165
	MLP43	At1g70890
	MLP168	At1g35310
	ExprePro	At3g08030
	CL12-A	At3g27830
	CL12-C	At3g27850

4.3.2 The PCR-Sequencing-Dot-blot strategy

Devising an efficient and economic way to identify the 774 clones was an important step in this project. PCR amplifying and subsequent sequencing the clones might be a quick way to find out the identities of the potential interactors, but PCR amplify all the 774 clones would consume a large

amount reagents and sequencing the PCR products would cost a lot of money. Since one clone may be represented many times among the 774 clones, if one clone is used as a probe in a dot-blot, all the identical clones will be fished out simultaneously and they would not require further PCR and sequencing. If the dot-blot is repeated using the identified clones as probes, the number of clones left to be PCR amplified and sequenced would be greatly decreased. Therefore, PCR-Sequencing and Dot-blot were combined to identify the potential interactors.

If the 27 probes were used individually in the dot-blots to figure out how frequently individual clones turned up, 27 dot-blots would have to be carried out for each of the 8 membranes (216 dot-blots in total), which would be very time consuming. Instead, we used pooled probes for the dot-blots so that the work can be reduced to 1/3 or less. For instance, we used 4 probes at the same time in the first round dot-blot. Instead of doing 32 dot-blots, we did 8 dot-blots only. In addition, after doing the dot-blot for several rounds, the number of unidentified clones reduced significantly. The unidentified clones were reloaded onto three 3 small membranes so that the work was reduced by another 75%.

All the dot-blots worked at the first attempt, and they gave strong and clear signals hence convincing results. The combination of PCR-Sequencing and Dot-blots was proved to be an economical and sensitive way to identify the potential interactors.

Chapter 5 Verification of the potential interactors in the RRS using the i1, i2, i1-i2 and i1-GGG-i2 baits

5.1 Introduction

Library screening using the yeast two-hybrid system is known to generate false positives (Hengen, 1997; Serebriiskii and Golemis, 2001; Causier, 2004), the number of which can be kept to a minimum by employing a stringent screening procedure (Lehner *et al*., 2004; Vidalain *et al*., 2004). In the RRS the ability to independently control bait and prey expression, and to screen for temperature revertants, enhance the stringency of the screen but are not infallible. It is also considered good practice to rescue and re-transform preys to confirm that interactions observed are due to those preys and not some other unidentified clone. For these reasons it is therefore necessary to verify the potential interactors to help eliminate false positives and to select the most promising candidates for further investigation. This step is especially important in this project because the bait used in the primary screen was a fusion of non-contiguous sequences.

As described in section 3.2.1, the i1 and i2 loops are so short that they were fused together and used as a single bait (i1-i2) in the initial library screening. Since the fusion of i1 and i2 may generate a potential binding site at the fusion point between i1 and i2. It is possible that some of the 20 potential interactors interacted with i1-i2 at the fusion point rather than the loop regions. It was hence necessary to create a bait which can be used to eliminate interactions based on the fusion point. The i1-GGG-i2 bait was designed with three glycine residues in between i1 and i2 to separate the two loops so that any interactions based on the artifact can be eliminated. Also, by spacing the loops out, they could be considered less likely to interfere with each other and disrupt interactions. If an interactor binds to i1-i2 but not i1-GGG-i2, it is likely that their interaction is based on the sequences at the boundary generated

between these non-contiguous sequences. On the contrary, if an interactor interacts with both i1-i2 and i1-GGG-i2, it is most likely to interact with either or both of i1 or i2. We can not eliminate the possibility that the three glycine residues might in some way interfere with an interaction between a prey and one or both of the loop regions and for this reason individual i1 and i2 baits can be used to trace whether each potential interactor binds to i1, i2, neither or both. In theory, a genuine interactor for i1-i2 should be expected to interact with at least one of the three additional baits.

5.1.1 Research objectives and experimental approach

20 potential interactors were identified in the initial library screening using the i1-i2 bait. In this chapter, we aimed to verify these potential interactors using the following approaches: (1) make the i1, i2 and i1-GGG-i2 baits for use in the RRS; (2) recover the prey plasmid that carries the sequence coding the potential interactor and re-transform it to the bait-containing yeast cells; (3) use the i1-i2 bait to reconfirm the interactions seen in the initial library screening; (4) use the i1-GGG-i2 bait to eliminate false positives that are based on fusion point of the i1-i2 bait; (5) use the i1 and i2 baits to trace whether each potential interactor interacts with i1, i2, neither or both.

5.2 Results

5.2.1 Making the pMetRas-i1, pMetRas-i2 and pMetRas-i1-GGG-i2 bait constructs

The inserts i1, i2 and i1-GGG-i2 were amplified by PCR using the Expand High Fidelity^{PLUS} PCR System (section 2.2.2.1, figure 5.1). The i1 and i2 inserts were amplified using the *GCR1* DNA as template, with the LP1FXmaI + LP1RXmaI and the LP2FXmaI + LP2RXmaI primer pairs respectively (Table 2.5). Two rounds of PCR were performed in order to make the i1-GGG-i2 insert. In the first round PCR, i1-GGG (GCR1 start codon to the end

of i1) and GGG-i2 (from the beginning of i2 to the end of i3) were amplified separately using the *GCR1* DNA as template, with GCR1bgn+LP1GGGrev and GGGLP2for+LP3RXmaI primer pairs respectively. In the second round PCR, i1-GGG and GGG-i2 were joined together for i1-GGG-i2 using the 1:100 dilutions of the first round i1-GGG and GGG-i2 PCR products as templates, with the LP1FXmaI + LP2RXmaI primer pair (Table 2.5). The predicted size of each insert is: i1: 54 bp, i2: 72 bp, i1-GGG-i2: 108 bp and the products obtained were consistent with this.

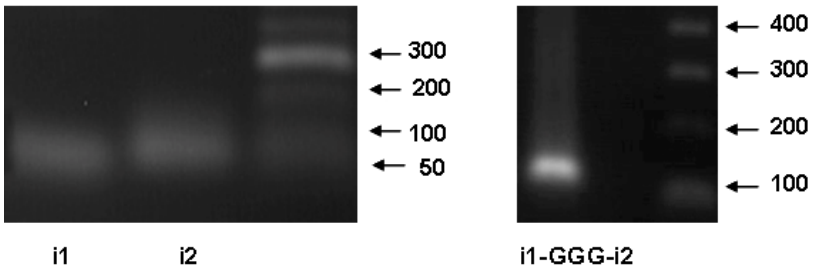


Figure 5.1. PCR amplification of the inserts. The three inserts i1, i2 and i1-GGG-i2 were PCR amplified. The predicted size of each insert is: i1: 54 bp, i2: 72 bp, i1-GGG-i2 108 bp.

PCR products were digested with *XmaI* and ligated into *XmaI* digested bait expression vector pMetRas. Recombinants were checked for insert size and orientation by colony PCR with a vector specific forward primer pMetF, and an insert-specific reverse primer LP1RXmaI or LP2RXmaI (Figure 5.2). Positive clones were identified, and the plasmids were extracted and sequenced (data not shown). Sequencing results confirmed that all inserts were successfully cloned in frame to the N-terminus of Myc which is translational fused to the N-terminus of Ras in pMetRas (*XmaI* site), which generated the pMetRas-i1, pMetRas-i2, and pMetRas-i1-GGG-i2 constructs.

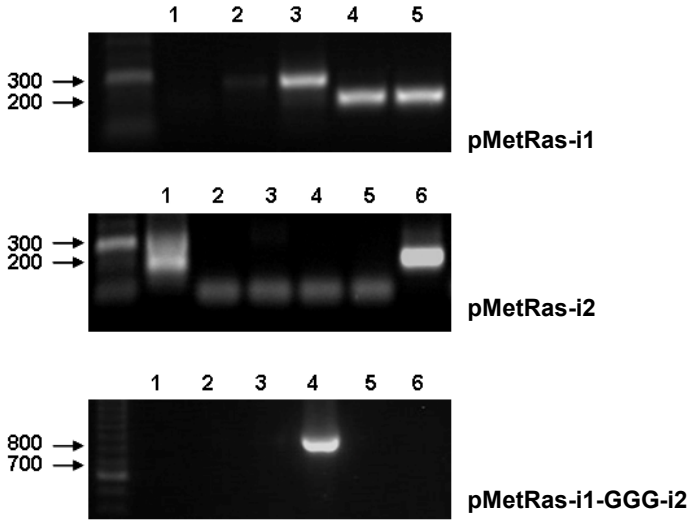


Figure 5.2 Colony PCR. Colony PCR was conducted with a vector specific forward primer pMetF, and an insert-specific reverse primer LP1RXmaI or LP2RXmaI to screen for positive clones. Positive clones were found for pMetRas-i1 (lane 4 and 5), pMetRas-i2 (lane 6) and pMetRas-i1-GGG-i2 (lane 4).

5.2.2 Detecting the expression of the i1, i2 and i1-GGG-i2 baits

The pMetRas-i1, pMetRas-i2 and pMetRas-i1-GGG-i2 constructs were transformed into the temperature sensitive yeast strain *cdc25-2*. Total proteins were extracted from the yeast cells that were grown in Glu-L-M media for bait expression and Glu-L+4M media for bait suppression. The pMetRas-i1-i2 construct was included in the experiment as a control. Because the bait vector contains a Myc epitope tag, the expression of the baits was checked by Western blot using the anti-Myc antibody. In the absence of methionine (-methionine), polypeptide of the expected size for the i1-Myc-Ras fusion was expressed at an extremely low level whereas polypeptides of the expected size for the i2-Myc-Ras, i1-GGG-i2-Myc-Ras and i1-i2-Myc-Ras fusions were expressed at a high level (Figure 5.3). In the presence of methionine (+methionine), no expression was detected for i1-Myc-Ras but very low level of expression was detected for i2-Myc-Ras, i1-GGG-i2-Myc-Ras and i1-i2-Myc-Ras (Figure 5.3).

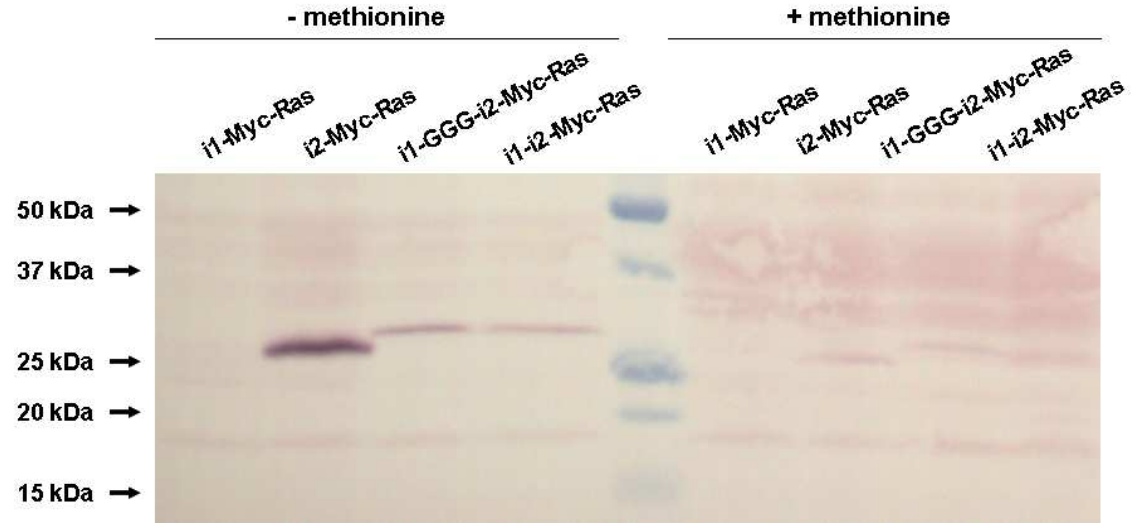


Figure 5.3 Western blot detection of bait-Myc-Ras fusion protein expression. Total proteins were extracted from the bait-containing (i1, i2, i1-GGG-i2, i1-i2) yeast cells that were grown in Glu-L-M media (-methionine) for bait expression and Glu-L+4M media (+methionine) for bait suppression. The expression of the baits was checked by Western blot with the anti-Myc antibody. The expected size of each bait fusion protein is: i1-Myc-Ras: 24.2 kDa, i2-Myc-Ras: 24.9 kDa, i1-GGG-i2-Myc-Ras: 26.63 kDa, and i1-i2-Myc-Ras: 26.0 kDa.

5.2.3 Autoactivation test of the i1, i2 and i1-GGG-i2 baits

Four yeast colonies for each of the three baits (i1, i2 and i1-GGG-i2) were randomly selected for the temperature sensitivity test. They were streaked out onto Glu-L+M plates and grown at 24°C for 3 days. These bait-containing yeast cells were then replica plated onto two Glu-L+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as control. Very little if any growth of replica plated cells was seen at 36°C compared with the extensive growth seen at 24°C on Glu-L+M (Figure 5.4). Therefore these bait-containing yeast cells were temperature sensitive and were used in subsequent experiments.

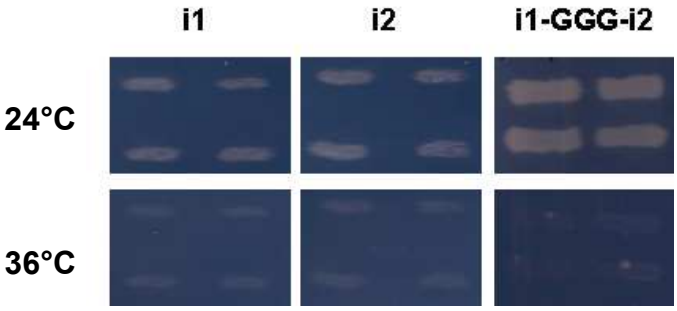


Figure 5.4 Bait temperature sensitivity test. The bait-containing (i1, i2 or i1-GGG-i2) yeast colonies were streaked onto Glu-L+M plate and grown at 24°C for 3 days. The yeast cells were then replica plated onto two Glu-L+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as a control. Photos show growth after this 3 day period.

As mentioned in section 3.2.4, it is vital to work only with baits that do not autoactivate the RRS. To check for autoactivation, yeast cells expressing bait with no prey were replica plated onto two Glu-L-M plates, one was grown at 36°C for 5 days, and the other at 24°C for 5 days as control. As shown in figure 5.5, virtually no growth of replica plated cells was seen at 36°C compared with the extensive growth seen at 24°C on Glu-L-M plates. The result implicated that none of the baits showed autoactivity, they are therefore suitable for use in the RRS.

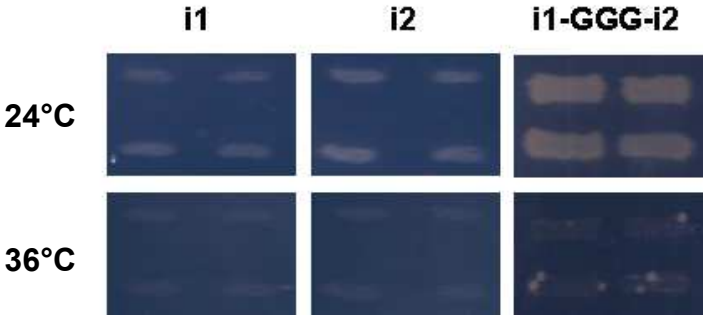
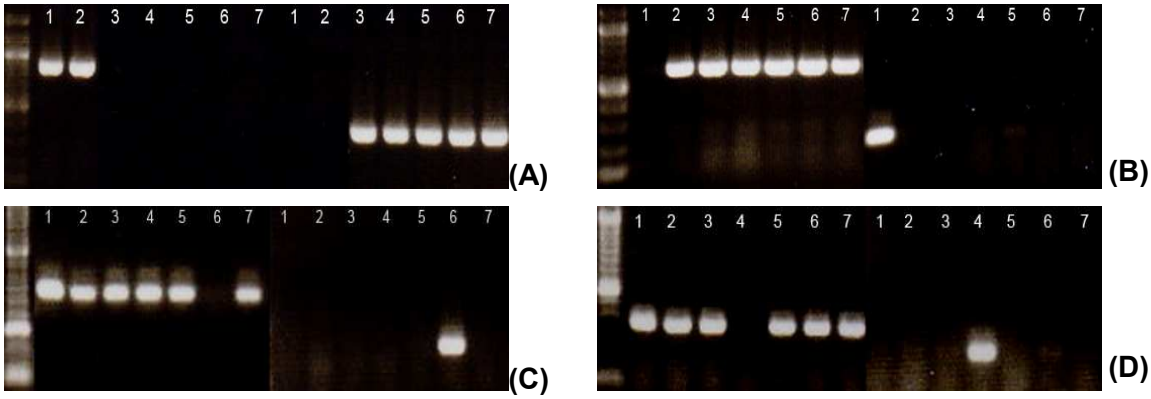


Figure 5.5 Bait autoactivation test. To check whether the baits autoactivate the RRS or not, the bait-containing yeast cells were replica plated from Glu-L+M plate onto two Glu-L-M plates. One was grown at 36°C for 5 days, and the other at 24 °C for 5 days as a control. Photos show growth after this 5 day period.

5.2.4 Recovery of the prey plasmids from yeast containing bait and potential interactors

The prey plasmid which carries the coding sequence of the potential interactor needed to be recovered so that it can be transformed into the bait-containing (i1, i2, i1-i2, i1-GGG-i2) yeast cells for subsequent experiments. Plasmids were extracted from yeast using the method described in section 2.2.12. Because the yeast from which the plasmids were isolated contains both bait and prey plasmids, and only the prey plasmid was wanted, the plasmid mixture was transformed into competent *E.coli* cells and plated out on LB-amp plates. Some of the colonies appearing on the plates would be expected to contain the bait plasmid alone, some would be expected to contain the prey plasmid alone, and others would contain both plasmids. Colony PCR was performed on randomly selected *E.coli* colonies to identify the ones that contain only the prey plasmid. Thus, each colony had undergone two rounds of colony PCR, the first using a prey plasmid specific forward primer T7 and a prey specific reverse primer, e.g. TRX2End, to check for prey plasmid presence, and the second using a bait plasmid specific forward primer pMetF and a bait specific reverse primer e.g. LP2RXmaI to check for bait plasmid absence (Figure 5.6). The *E.coli* colonies that contain the prey plasmid alone were patched out on LB-amp plates and the plasmid was extracted from the cells that grew after 24 hours.



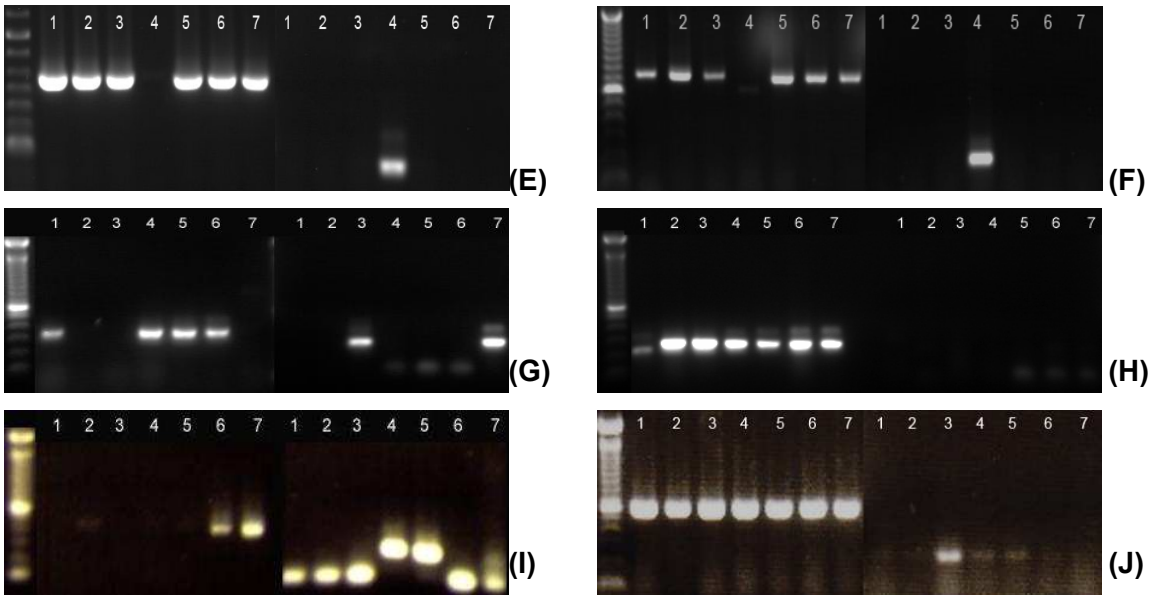


Figure 5.6 Examples of *E.coli* colony PCR screening for colonies that contain the prey plasmid alone. For each potential interactor, 7 randomly picked E. coli colonies had undergone two rounds of PCR. The 7 lanes on the left were the PCR products of colony No.1-7 amplified using T7 and a gene specific reverse primer to check prey plasmid presence. The 7 lanes on the right were PCR products of colony No.1-7 amplified using the pMetF and LP2RXmaI primers to check bait plasmid absence. Examples shown are (A)TRX2, (B)ZAC, (C)ASK1, (D)AHB2, (E)*zf*-DHHC1, (F)C2dcp, (G)QPRTase, (H)ADPGlc PPase, (I)ExprePro, (J)MLP168.

While it was relatively easy to isolate colonies containing the prey plasmid alone for some potential interactors, for others, such TRX2 and ExprePro, it was more difficult and I had to screen more than 20 colonies to recover these prey plasmids. For certain interactors, such as PRMT and cytochrome P450, I could not recover them. This could be because the *E.coli* strain was unable to support the replication of some of the prey DNAs, or because prey proteins were expressed at a low level and were toxic to *E.coli*. I was able to recover the prey plasmids for the following potential interactors: TRX2, TRX3, TRX4, AHB2, *zf*-DHHC1, C2dcp, ZAC, LRR RLK, ASK1, PAG1, QPRTase, ADPGlc PPase, ExprePro, and MLP168.

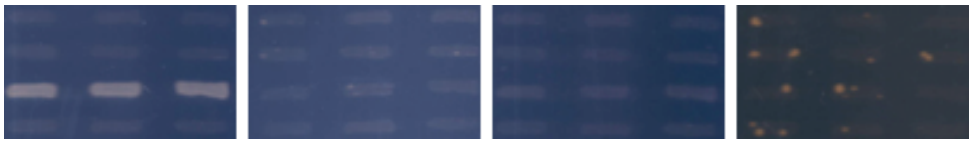
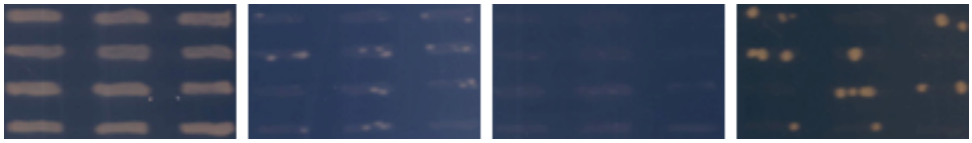
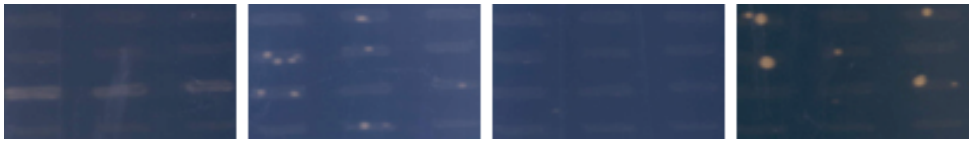
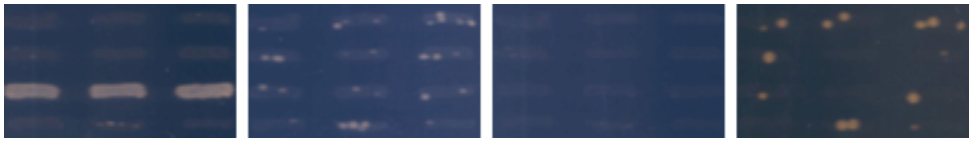
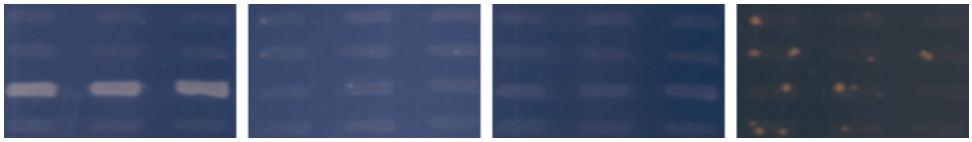
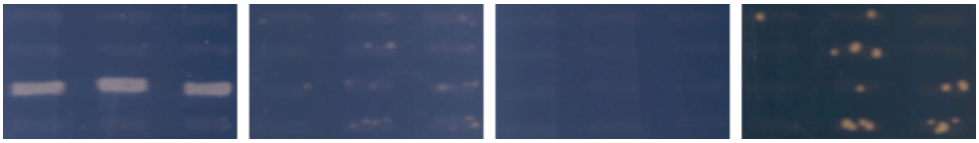
5.2.5 Verification of the potential interactors using the i1, i2 i1-i2 and i1-GGG-i2 baits

The recovered prey plasmids were transformed into the bait-containing (i1, i2, i1-i2, i1-GGG-i2) yeast cells and plated onto Glu-L-U+M plates and incubated at 24°C until colonies appeared. The colonies were streaked onto Glu-L-U+M plate in the format shown in table 5.1, and were incubated at 24°C until they had grown sufficiently to be replica plated. It should be noted that there were three streaks for each row, which contain the same plasmids but were taken from different transformants. This is to ensure that results seen are reproducible.

Table 5.1 Yeast streaks arranged on the Glu-L-U+M plate for replica plating

i1 + prey	Streak 1	Streak 2	Streak 3
i2 + prey	Streak 1	Streak 2	Streak 3
i1-i2 + prey	Streak 1	Streak 2	Streak 3
i1-GGG-i2 + prey	Streak 1	Streak 2	Streak 3

The yeast streaks were replica plated from the Glu-L-U+M plate onto the selective medium Gal-L-U-M and the three control media Glu-L-U-M, Gal-L-U+4M and YPD, and were incubated at 36°C for 7 days. The plates were scanned every 24 hours from day 3 until day 7. Figure 5.7 shows the plates that were scanned on day 7.

Media / Expression / Prey	Gal-L-U-M Bait + Prey	Glu-L-U-M Bait	Gal-L-U+4M Prey	YPD None
TRX2				
TRX3				
TRX4				
AHB2				
zf-DHHC1				
C2dcp				
ZAC				
LRR RLK				
ASK1				

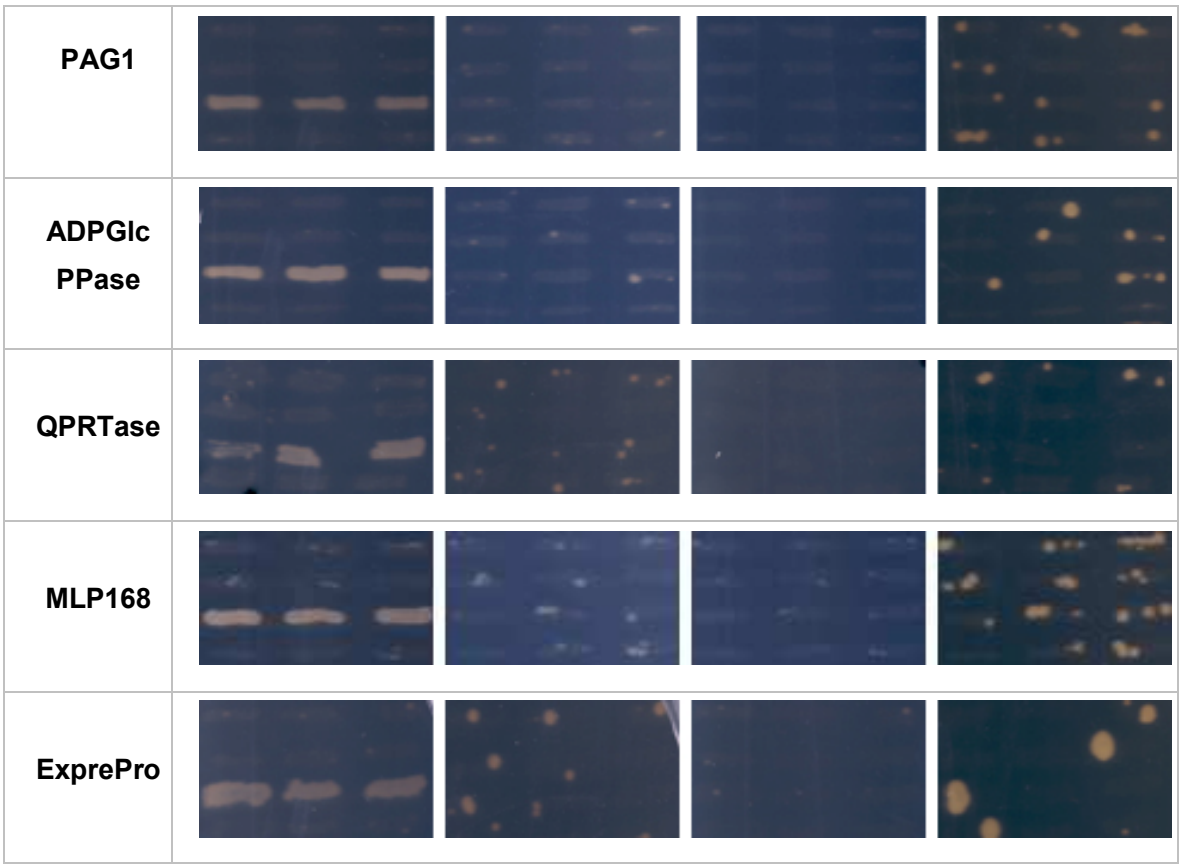


Figure 5.7 Interactions between potential interactor (prey) and the i1, i2, i1-i2 and i1-GGG-i2 baits. Yeast streaks were replica plated from the Glu-L-U+M plate onto the selective medium Gal-L-U-M and the control media Glu-L-U-M, Gal-L-U+4M and YPD, and were incubated at 36°C for 7 days. Photos show growth after this 7 day period.

As shown in figure 5.7, all potential interactors except LRR RLK interacted with the i1-i2 bait. However, only TRX3, TRX4 and *zf*-DHHC1 interacted with the i1-GGG-i2 bait. Again, only TRX3, TRX4 and *zf*-DHHC1 interacted with both the i1 and the i2 bait. The result was consistent across the three streaks for each bait + prey combination.

5.3 Discussion

The i1-i2 bait which was original used in the library screening, together with the three new baits i1, i2, and i1-GGG-i2 were used to further verify the potential interactors. The expression of the bait fusion proteins was checked by Western blot using the anti-Myc antibody. In the absence of methionine,

the i2-Myc-Ras, i1-GGG-i2-Myc-Ras and i1-i2-Myc-Ras fusions were expressed at a high level whereas i1-Myc-Ras was expressed at an extremely low level such that it was barely visible on the PVDF membrane. Nevertheless, three interactors showed interaction with the i1 bait (Figure 5.7, TRX3, TRX4 and *zf*-DHHC1), which confirmed that i1-Myc-Ras bait fusion was indeed expressed. In the presence of methionine, no expression was detected for i1-Myc-Ras, but a low level of expression was detected for i2-Myc-Ras, i1-GGG-i2-Myc-Ras and i1-i2-Myc-Ras. This phenomenon that some baits express at a low level in the presence of methionine has been observed before (section 3.3.2), and the possible reason has been discussed, so it will not be repeated here.

The i1-i2 bait is artificial because it is a fusion between adjacent but non-contiguous amino acid sequences. As such it would not occur in the native state of the receptor. Thus some interactions with it may occur because the fusion bait does not represent the native state. To decipher whether this was the case, the new bait i1-GGG-i2 was included which had introduced three glycine residues between i1 and i2. This should introduce a spatial gap between the two loops and allow detection of interactions that may have been centered on this non-native region generated at the fusion point between i1 and i2. Most of the potential interactors did not interact with this bait suggesting that their interaction with i1-i2 might not be a feature of the native protein (Figure 5.7). It is likely that these interactions occur at the artifact of the fusion point, however we can not rule out the possibility that the three glycine residues somehow disrupt a bona fide interaction between the preys and one or more of the loop regions. The inclusion of individual loops i1 and i2 help to examine this question in more detail. The evidence supporting this hypothesis was that these potential interactors were also unable to interact with i1 and i2, which indicates that they do not bind to i1 and i2, therefore their interactions with i1-i2 might be false positives. However, we can not exclude the possibility that they are genuine interactors but did not show interaction due to some other factor such as inappropriate physiological conditions, such

as pH and temperature. Lack of time prohibited performing more detailed analyses on these interactors.

TRX3, TRX4 and *zf*-DHHC1 distinguished themselves from all other potential interactors by interacting with all the baits, and are therefore considered most likely genuine interactors. They were certainly the most promising candidates, and were therefore chosen for further investigation. Judging from the results obtained, we conclude that the verification work was very successful and was a very important step in this project.

Chapter 6 Characterisation of the GCR1-TRX3 and GCR1-TRX4 interactions

6.1 Introduction

Two members of the *Arabidopsis* thioredoxin *h* family – TRX3 and TRX4 were identified as potential GCR1 interactors in the initial library screening (chapter 5). This chapter describes the work carried out on the characterisation of the GCR1-TRX3 and GCR1-TRX4 interactions.

6.1.1 Thioredoxins

Thioredoxins are small proteins (12 -14 kDa) capable of catalyzing thiol-disulfide redox reactions (Holmgren 1995; Peterson *et al*., 2005). They are present in all organisms from prokaryotes to eukaryotes, and are characterized by the conserved active site WCXPC (X=G or P) which contains two cysteines (Gelhaye *et al*., 2004a; Kumar *et al*, 2004). They reduce disulfide bridges of target proteins by the reversible formation of a disulfide bridge between two neighbouring cysteine residues present in the active site. Thioredoxins are maintained reduced via NADPH-dependent thioredoxin reductase (NTR) in cytosol and mitochondria, or via ferredoxin-dependent thioredoxin reductase (FTR) in the chloroplast or cyanobacteria (Rouhier and Jacquot, 2003; Reichheld *et al*., 2005).

Two thioredoxins have been identified in *E.coli* (reviewed in Carmel-Harel and Storz, 2000). Two cytoplasmic thioredoxins and a mitochondrial thioredoxin have been identified in *Saccharomyces cerevisiae* (Pedrajas *et al*., 1999; reviewed in Trotter and Grant, 2005). The human genome contains one cytoplasmic thioredoxin and one mitochondrial thioredoxin (Wollman *et al*., 1988; Damdimopoulos *et al*., 2002). In contrast to other organisms, plants have a particularly complex thioredoxin system. Approximately 40 thioredoxin and

thioredoxin-related proteins have been identified (Meyer *et al*., 2006). They are divided into two families based on sequence similarity. Family I include proteins that contain one distinct thioredoxin domain and it can be further divided into six groups – cytosolic thioredoxins *h*, chloroplastic thioredoxins *f, m, x* and *y,* and mitochondrial thioredoxins *o*. Family II is composed of fusion proteins with one or more thioredoxin domains coupled to additional domains (Gelhaye *et al*., 2004a). The thioredoxins TRX3 and TRX4, which we identified as potential GCR1 interactors, belong to thioredoxins *h*. The fact that they are predicted to be in the cytoplasm is consistent with the location of the regions of GCR1 used as baits. Eight members of the thioredoxin *h* group have been identified in *Arabidopsis*. They can be divided into three subgroups according to the analysis of their primary structure (Figure 6.1). AtTRX*h*1, *h*3 (TRX3), *h*4 (TRX4) and *h*5, which are approximately the same size as *E. coli* thioredoxin, belong to subgroup 1. AtTRX*h*2, *h*7 and *h*8 belong to subgroup 2, and *h*9 belongs to subgroup 3 (Gelhaye *et al*., 2004a). Both subgroup 2 and 3 contain an N-terminal extension with no established function (Brehelin *et al*., 2004).

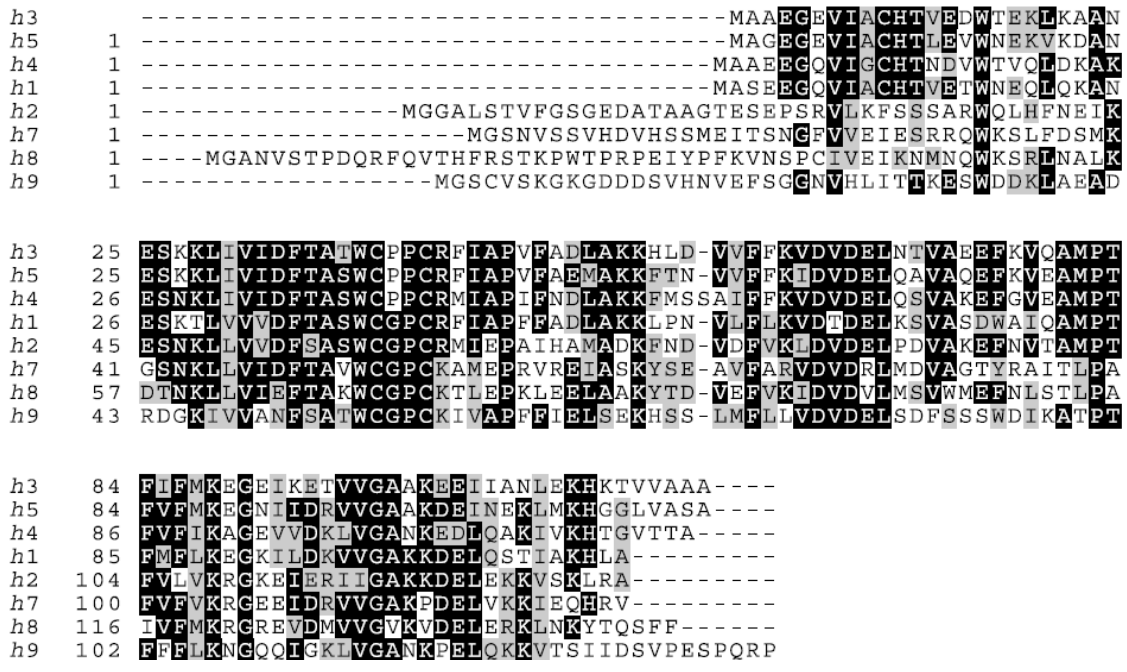


Figure 6.1 Multiple sequence alignment of the *Arabidopsis* thioredoxin *h* group (taken from Brehelin *et al*., 2004).

Thioredoxin was originally identified as an electron donor for reductive enzymes, such as ribonucleotide reductase, thioredoxin peroxidases and methionine sulfoxide reductases (Arner and Holmgren, 2000). However, the reducing capability of thioredoxins allows them to interact with a large number of macromolecules and play roles in many cellular processes. They can act as powerful antioxidants by reducing reactive oxygen species (ROS), e.g. hydrogen peroxide (H_2O_2), via the enzyme thioredoxin peroxidase (Saitoh *et al*., 1998). They can also modify the activity of various receptors via disulphide bond formation. For instance, human mitochondrial thioredoxin-2 blocks tumor necrosis factor-α (TNF-α)-induced ROS generation and inhibits subsequent signalling events, including transcription factor NF-κB activation and apoptosis (Hansen *et al*., 2006). The functional activity of the steroid hormone glucocorticoid receptor (GR) is suppressed under oxidative conditions, while overexpression of thioredoxin counteracts this (Makino *et al*., 1996; Okamoto *et al*., 1999). Further investigation revealed that direct association between the thioredoxin and the DNA binding-domain of GR allows redox regulation of GR function, indicating a critical role of thioredoxin in modulating steroid receptor-mediated signal transduction (Makino *et al*., 1999). Thioredoxins are also involved in the regulation of enzyme activities. *E.coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7 (Tabor *et al*., 1987). Mammalian thioredoxin can bind to and inhibit the apoptosis signal-regulating kinase-1 (ASK-1) activity and the subsequent ASK-1-dependent apoptosis (Saitoh *et al*., 1998). In recent years, a growing body of evidence indicates that thioredoxins can also act as protein chaperones, facilitating protein folding independent of their role as reducing catalysts (Kern *et al*., 2003; Jurado *et al*., 2006). For instance, Jurado and co-workers (2006) have demonstrated that thioredoxin fusions increase the correct folding of single chain Fv antibodies in the cytoplasm of *E.coli*.

In plants, thioredoxins play fundamental roles in a wide variety processes. They participate in the mobilization of protein reserves during seed germination and play a central role in the redox conversion of seed proteins (Basse and

Buchanan, 1997; Alkhalfioui *et al*., 2007). In the legume *Medicago* the G-protein β subunit has been detected as a thioredoxin binding protein and was found to be partly reduced in the axis and cotyledons of dry seeds and further reduced during germination (Wong *et al*., 2004; Alkhalfioui *et al*., 2007). Thioredoxins can also act as scavengers of ROS and as components of signalling pathways in plant response to oxidative stress (Santos and Rey, 2006). Certain plants can recognise and reject self-pollen to prevent self-fertilization. In *Brassica*, two thioredoxins, THL1 and THL2, have been shown to interact with the kinase domain of the S Receptor Kinase (SRK) which is the female determinant of self-incompatibility (Bower *et al*., 1996; Mazzurco, *et al* 2001). They have been suggested to function as negative regulators of SRK to reduce its basal activity (Mazzurco, *et al* 2001). The chloroplastic thioredoxins participate in the regulation of enzymatic activities during the transition between light and dark phases (Brehelin *et al*., 2004), and are involved in cell division and plant reproduction (Barajas-Lopez *et al*., 2007).

The structure of thioredoxins has been determined by x-ray crystallography and NMR from various organisms, in both reduced and oxidized forms (Jeng *et al*., 1994; Qin *et al*., 1994; Holmgren, 1995; Brehelin *et al*., 2000). Overall, the structures of the reduced and oxidized states of thioredoxin are very similar and the packing of side chains within the protein core is nearly identical (Qin *et al*., 1994). It consists of a central core of five β-sheets enclosed by four α-helices. The WCXPC active site is located on the surface of the protein, at the end of a β-strand and at the beginning of a long α-helix (Kern *et al*., 2003). The interaction between thioredoxin and substrate proteins has been suggested to involve the active site and several residues such as Pro76 and Gly92 (numbered according to *E.coli* thioredoxin 1), which form a moderately hydrophobic surface around the active site and facilitate interactions with other enzymes (Kern *et al*., 2003). The hydrophobic area around the active site has also been shown to be involved in determining target specificity (Holmgren 1995).

6.1.2 The reverse Ras-Recruitment System (rRRS)

The rRRS is a reverse version of the RRS. It can be used as a simple and quick method to confirm protein-protein interaction seen in the RRS. The idea is to alter the localisation of the bait and prey proteins. So a prey protein that was originally localised to the membrane by the myristoylation sequence would instead be attached to the mRas and localised in the cytoplasm. Likewise, the former bait protein would no longer be attached to the mRas but the myristoylation sequence if required, and localised to the cell membrane. When the prey couples to the bait, the mRas will be brought to the cell membrane and complement the mutant yRas, thus enabling yeast growth at the restrictive temperature.

The advantage of using the rRRS to reconfirm the interaction seen in the RRS is that, instead of using discrete portions of GCR1, the full length receptor can be used and can be expected to localise to the membrane in its native conformation. It overrides the problem of the prey recognizing the possible non-native binding domain(s) in the RRS. If the interactions first seen in the RRS occur in the rRRS, then it is more likely to be a genuine protein-protein interaction. The rRRS has been used successfully to identify novel interactors of membrane proteins that are used as baits in their native conformation (Hubsman *et al*., 2001; Frankel *et al*., 2005), hence serving as a good alternative for investigating protein-protein interactions of membrane proteins.

6.1.3 Affinity chromatography (pull-down)

Although the rRRS can be used as a good approach to reconfirm the interactions seen in the RRS, it is prudent to verify the interactions using an independent method, such as affinity chromatography (pull-down, section 1.2.1.1). Briefly, in a pull-down assay, a tagged bait protein is captured on affinity beads that are specific for the tag, and incubated with a pool of proteins which contain the prey protein. After the washing step which eliminates most

non-specific binding, the bait together with the prey that binds to it are eluted and the elution fraction is further analysed.

A stable protein-protein interaction should be relatively easy to isolate by a pull-down assay because it should be able to persist the washing steps. In contrast, transient and weak interactions are likely to be more difficult to identify through pull-downs, because they may disassociate over time or during the washing steps. However, optimizing assay conditions, such as pH, salt species and salt concentration, as well as carrying out appropriate control experiments may help to obtain ideal results. In recent years, a very large number of studies have demonstrated that the pull-down technique has become an invaluable tool for the verification of protein-protein interactions predicted by other research methods, e.g. yeast two-hybrid and co-immunoprecipitation. For this reason, besides the rRRS, the pull-down method was also used in the project to reconfirm the interactions seen in the RRS.

6.1.4 Research objectives and experimental approach

Thioredoxins usually interact with a protein via cysteine or methionine residues. However none of the i1 or i2 baits has these features. Interestingly, there are two cycteines (-VLCYCLF-) in the TM1 domain that are very close to i1 (Figure 3.1). As described in section 3.3.1, the boundaries of the intracellular domains may move up and down in relation to the membrane when the protein is in different conformations. Therefore these two cysteines could possibly be exposed to the cytoplasm under certain circumstances, becoming a potential binding site for the target proteins. We were interested in whether TRX3 and TRX4 interact with the extended i1 bait containing part of the TM1 domain with these two cysteines (VLCYCLF-i1 and CYCLF-i1). Therefore, the first objective of this chapter was to test whether TRX3 and TRX4 interact with the VLCYCLF-i1 and CYCLF-i1 baits, and whether the strengths of their interactions are different compared with the i1 bait. It is worth mentioning that besides the CYCLF-i1 bait which contains the two cysteine residues, the VLCYCLF-i1 bait

was also included in the test for the reason that the addition of the VL residues to CYCLF-i1 would put the first cysteine residue in an environment that is more similar to its native contact.

Previous studies have shown that substitution of the second cysteine in the active site of thioredoxin by a structural analog, such as serine, enables thioredoxin to form a stable intermediate complex with the target protein which contains cysteines (Nishiyama *et al*., 1999; Yamanaka *et al*., 2000; Broin *et al*., 2002). If TRX3 or TRX4 could reduce the cysteines in TM1 of GCR1, mutant TRX3 or TRX4 that contain a CPPC to CPPS active site mutation are expected to form an intermediate complex with VLCYCLF-i1 or CYCLF-i1, and thus prolong or enhance the interaction. This could be reflected by an increased level of yeast growth in the RRS. Therefore, we constructed the active site mutants S42 (for TRX3) and S43 (for TRX4) by substitute the second cysteine with serine (Figure 6.2). The S39, S39S42 (for TRX3) and S40, S40S43 (for TRX4) mutants were also constructed and used as controls (Figure 6.2). The mutants together with wild type TRX3 and TRX4 were used as preys to test their interaction with the VLCYCLF-i1, CYCLF-i1, i1, i2, i1-i2 and i1-GGG-i2 baits in the RRS.

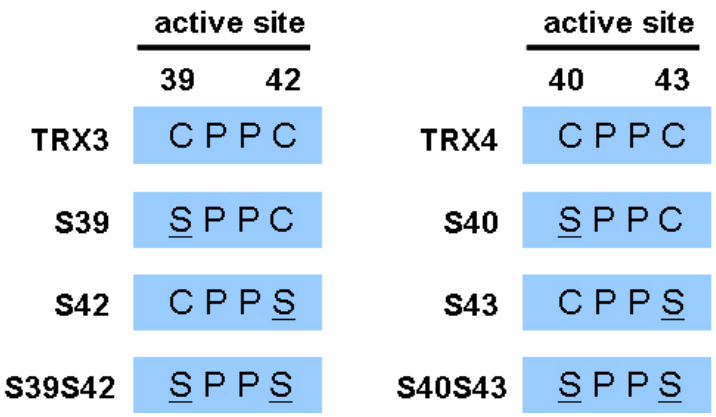


Figure 6.2 Schematic representation of mutant (S39, S42, S39S42, S40, S43 and S40S43) and wild type TRX3 and TRX4. C: Cysteine; P: Proline; S: Serine. 39, 42, 40 and 43 represent the amino acid number in the corresponding protein sequence.

The third objective was to examine whether TRX3 and TRX4 interact with other parts of GCR1 besides the i1 and i2 regions, by performing the rRRS screens using various parts of GCR1 as baits (from the beginning of Nter to the end of i2 (Nter-i2), from the beginning of i3 to Cter (i3-Cter), and the full length GCR1 (Nter-Cter)), and TRX3 and TRX4 as preys.

The last objective was to verify the interactions between GCR1 and TRX3 or TRX4 using pull-down assays. Since GCR1 is a membrane protein which contains seven transmembrane domains, it is very difficult to get its full length soluble protein in *E.coli*. The i1-GGG-i2 bait has been demonstrated to be the best bait to verify potential interactors in the RRS (section 5.3), it was deemed a suitable alternative to the full length GCR1 for use in the pull-downs. Therefore, pull-downs were performed using i1-GGG-i2 as bait, with TRX3 and TRX4 as preys.

6.2 Results

6.2.1 TRX3 and TRX4 interact with the extended i1 bait containing part of the TM1 domain with two cysteines (VLCYCLF-i1 and CYCLF-i1)

This section describes the result obtained for the first objective, which was to test whether TRX3 and TRX4 interact with the VLCYCLF-i1 and CYCLF-i1 baits, and whether the strengths of their interactions would change compared with the i1 bait.

6.2.1.1 Making the pMetRas-VLCYCLF-i1 and pMetRas-CYCLF-i1 bait constructs

The VLCYCLF-i1 and CYCLF-i1 inserts were PCR amplified using the Expand High Fidelity^{PLUS} PCR System (section 2.2.2.1, figure 6.3). VLCYCLF-i1 was amplified using pMetRas-i1 as template with the VLP1FXmaI + LP1RXmaI

primer pair. CYCLF-i1 was amplified using pMetRas-i1 as template, with the CLP1FXmaI + LP1RXmaI primer pair (Table 2.5). The predicted size of each PCR product is VLCYCLF-i1: 72 bp and CYCLF-i1: 69 bp, and the products obtained were consistent with this.

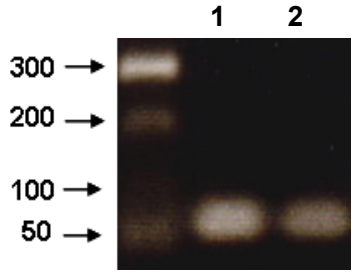


Figure 6.3 PCR amplification of the inserts. The VLCYCLF-i1 (lane1) and CYCLF-i1 (lane 2) inserts were PCR amplified. The predicted size of each insert is: VLCYCLF-i1: 72 bp, CYCLFi1: 69 bp.

The PCR products were digested with *XmaI* and ligated into *XmaI* digested RRS bait expression vector pMetRas. Recombinants were checked for insert size and orientation by colony PCR using a vector-specific pMetF as forward primer, and an insert-specific LP1RXmaI as reverse primer (data not shown). Positive clones were identified and the plasmids were extracted and sequenced (data not shown). Sequencing results confirmed that both inserts were successfully cloned in frame to the N-terminus of Myc which is translational fused to the N-terminus of Ras in pMetRas (*XmaI* site), which generated the pMetRas-VLCYCLF-i1 and pMetRas-CYCLF-i1 constructs.

6.2.1.2 Detecting the expression of the VLCYCLF-i1 and CYCLF-i1 baits

The pMetRas-VLCYCLF-i1 and pMetRas-CYCLF-i1 constructs were transformed into the temperature sensitive yeast strain *cdc25-2*. Total proteins were extracted from the yeast cells that were grown in Glu-L-M media for bait expression and Glu-L+4M media for bait suppression. Because the bait vector contains a Myc epitope tag, the expression of the bait fusion proteins was checked by Western blot using the anti-Myc antibody. The i1, i2, i1-i2, i1-GGG-i2 baits were included in the experiment as controls. In the absence of

methionine (Figure 6.4A), the VLCYCLF-i1-Myc-Ras fusion protein was either not expressed or was expressed at an extremely low level that could not be detected by this method. A polypeptide of the expected size (24.87 kDa) for the CYCLF-i1-Myc-Ras fusion was expressed at a very low level. A polypeptide of the expected size (24.2 kDa) for the i1-Myc-Ras fusion was expressed at a low level, but higher than the CYCLF-i1-Myc-Ras fusion. A polypeptide of the expected size (24.9 kDa) for the i2-Myc-Ras fusion was expressed at a very high level. A polypeptide of the expected size for the i1-i2-Myc-Ras fusion (26 kDa) was expressed at a level that is lower than the i2-Myc-Ras fusion but similar to the i1-GGG-i2-Myc-Ras fusion (26.63 kDa). A polypeptide of the expected size (23.05 kDa) for the Myc-Ras fusion encoded by the empty bait vector was also detected. In the presence of methionine (Figure 6.4B), no polypeptides of the expected size for the fusion proteins were detected except for i2-Myc-Ras and Myc-Ras. Overall, the order of the expression levels for the bait fusion proteins can be summarised as:

i2-Myc-Ras > i1-i2-Myc-Ras = i1-GGG-i2-Myc-Ras > i1-Myc-Ras > CYCLF-i1-Myc-Ras > VLCYCLF-i1-Myc-Ras

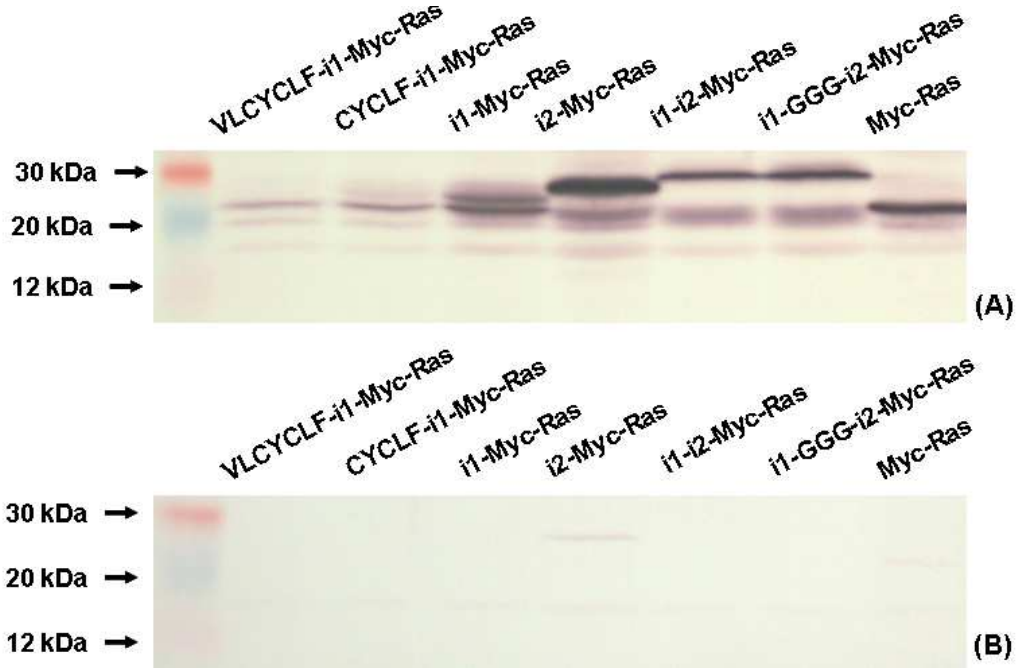


Figure 6.4 Western blot detection of bait-Myc-Ras fusion protein expression. Total proteins were extracted from yeast cells that were grown in Glu-L-M media for bait

expression (A) and Glu-L+4M media for bait suppression (B). The expression of the bait fusion proteins was checked by Western blot using the anti-Myc antibody. The expected size of each fusion protein is: VLCYCLF-i1-Myc-Ras: 25.01kDa, CYCLF-i1-Myc-Ras: 24.87kDa, i1-Myc-Ras: 24.20 kDa, i2-Myc-Ras: 24.9 kDa, i1-i2-Myc-Ras: 26 kDa, i1-GGG-i2-Myc-Ras 26.63 kDa, Myc-Ras: 23.05 kDa.

6.2.1.3 Autoactivation test of the VLCYCLF-i1 and CYCLF-i1 baits

Four yeast colonies for each of the two baits (VLCYCLF-i1 and CYCLF-i1) were randomly selected for temperature sensitivity test. They were streaked onto Glu-L+M plates and grown at 24°C for 3 days. These bait-containing yeast cells were then replica plated onto two Glu-L+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as control. There was virtually no yeast growth at 36°C, but all yeast streaks grew at 24°C (Figure 6.5). Therefore these bait-containing yeast cells were temperature sensitive, and were used in the following experiments.

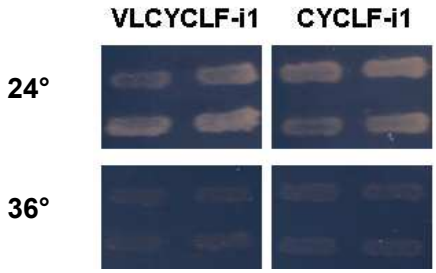


Figure 6.5 Bait temperature sensitivity test. The bait-containing (VLCYCLF-i1 and CYCLF-i1) yeast cells were streaked onto Glu-L+M plate and were grown at 24°C for 3 days. They were then replica plated onto two Glu-L+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as a control. Photos show yeast growth after this 3 day period.

It is vital to work only with baits that do not autoactivate the RRS. To check for autoactivation, yeast cells expressing bait with no prey were replica plated from a Glu-L+M plate onto two Glu-L-M plates, one was grown at 36°C for 5 days, and the other at 24°C for 5 days as a control. There was virtually no yeast growth at 36°C, but all yeast streaks grew at 24°C (Figure 6.6). The result indicated that the VLCYCLFi1 and CYCLF-i1 baits did not autoactivate RRS, therefore they were suitable for use in the subsequent RRS screens.

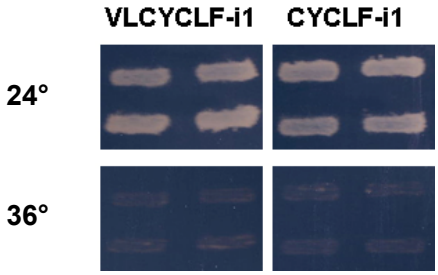


Figure 6.6 Bait autoactivation test. To check whether the baits autoactivate the RRS or not, the yeast cells were replica plated from a Glu-L+M plate onto two Glu-L-M plates, one was grown at 36°C for 5 days, and the other at 24°C for 5 days as a control. Photos show growth after this 5 day period.

6.2.1.4 TRX3 and TRX4 interact with VLCYCLF-i1 and CYCLF-i1 in the RRS

The pUra-TRX3, pUra-TRX4 as well as the pUra-TRX2 construct, which was used as a control, were transformed into the bait-containing (VLCYCLF-1 and CYCLF-i1) yeast cells and were plated onto Glu-L-U+M medium and incubated at 24°C until colonies appeared. The colonies were streaked onto the Glu-L-U+M plates and incubated at 24°C until they had grown sufficiently to be replicate plated. They were then replica plated onto the selective medium Gal-L-U-M and the three control media Glu-L-U-M, Gal-L-U+4M and YPD, and incubated at 36°C for 6 days. The interactions for the i1, i2, i1-i2 and i1-GGG-i2 baits with the TRX2, TRX3 and TRX4 preys were included in the experiment as controls. The plates were scanned every 24 hours from day 3 until day 6 (Figure 6.7).

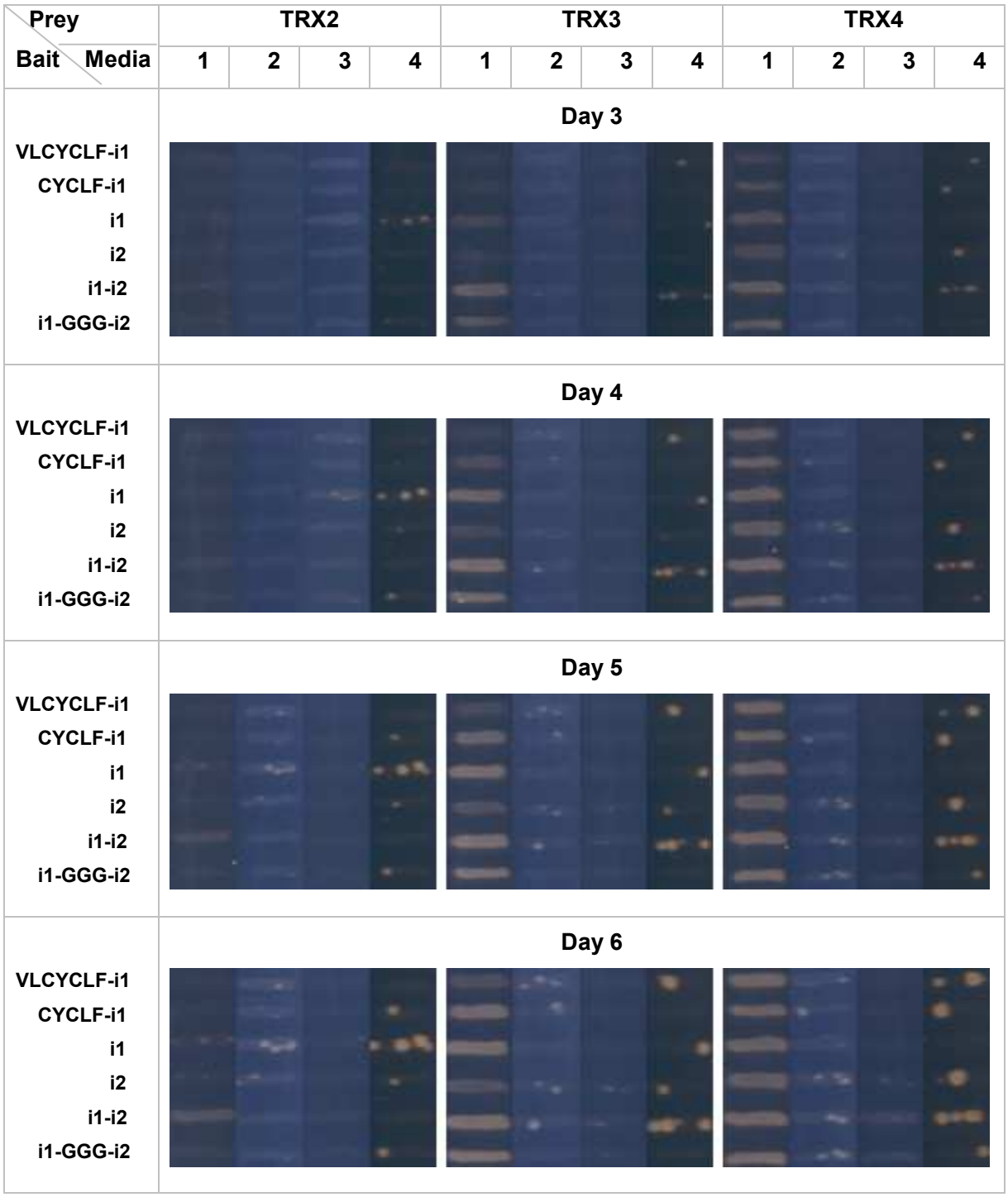


Figure 6.7 Interactions between GCR1 derived baits and the TRX2, TRX3 and TRX4 preys. Medium 1: Gal-L-U-M, 2: Glu-L-U-M, 3: Gal-L-U+4M, 4: YPD. The yeast streaks were replica plated onto the four media/plates and incubated at 36°C for 6 days. The plates were scanned every 24 hours from day 3 to day 6. It should be noted that the four plates were scanned simultaneously on a flatbed scanner to produce a single image file. Appropriate regions of the image has been selected and presented in this figure such that side by side comparison can be made. The differences in the colour of the background are due to different media compositions and colour of the individual plate from which the scan was taken.

As shown in figure 6.7, TRX2 did not interact with any of the GCR1 derived baits except i1-i2, which was the original bait that fished it out in the initial library screening. This observation was consistent with our previous result that TRX2 does not interact with i1 and i2 (section 5.3). On the contrary, both TRX3 and TRX4 interacted with all the baits, although the strength of interactions varies. This indicates that they both interact with not only the i1 and i2 baits, but also the extended i1 baits: VLCYCLF-i1 and CYCLF-i1.

For TRX3 (Figure 6.7), on the Gal-L-U-M plate, where both bait and prey were expressed, it can be seen that on day 3, there was much less growth for yeast containing VLCYCLF-i1, CYCLF-i1, i1 and i2 than yeast containing i1-i2 and i1-GGG-i2. The growth for i1 caught up with i1-i2 and i1-GGG-i2 on day 4. By day 5, the level of growth for CYCLF-i1 was similar to i1. However on day 6, the level of growth of VLCYCLF-i1 and i2 were still much lower than the others. Overall, the yeast containing i1-i2 and i1-GGG-i2 grew a lot faster than the others.

TRX4 exhibited a similar growth profile as TRX3 – yeast containing the i1-i2 and i1-GGG-i2 baits grew a lot faster than the others, followed by i1 and CYCLF-i1 as shown clearly on day 3 (Figure 6.7). However, the level of growth for yeast cells containing VLCYCLF-i1 and i2 caught up with i1 and CYCLF-i1 on day 5. Clearly, on day 6 the level of growth was almost identical for yeast cells containing the six individual baits. Interestingly, the growth for the TRX4 appeared to occur slightly earlier than TRX3.

6.2.2 GCR1 derived baits might not interact with the S39, S42, S39S42, S40, S43 and S40S43 preys in the RRS

This section describes the results obtained for the second objective, which was to investigate whether the reductive capability of the active site CPPC of TRX3 and TRX4 is required for their interaction with GCR1. It was achieved by testing

the interactions between the GCR1 derived baits and the TRX3 and TRX4 derived active site mutant preys in the RRS.

6.2.2.1 Making the pUra-S39, pUra-S42, pUra-S39S42, pUra-S40, pUra-S43 and pUra-S40S43 prey constructs for use in the RRS

All inserts (S39, S42, S39S42, S40, S43 and S40S43) were amplified by PCR using the Expand High Fidelity^{PLUS} PCR System (section 2.2.2.1, figure 6.8). Two rounds of PCR were performed in order to make the S39 insert. In the first round PCR, S39N (TRX3 aa 1-43) and S39C (TRX3 aa 36-119) were amplified separately using the *TRX3* cDNA as template, with the TRX3BegBmRI+TRX3^{Ser39}R and the TRX3^{Ser39}F+TRX3EndRI primer pairs respectively. In the second round PCR, S39N and S39C were joined together for S39 using the 1:100 diluted first round PCR products as templates, with the TRX3BegBmRI + TRX3EndRI primer pair (Table 2.5). The S42, S40 and S43 inserts were made the in same way as the S39 insert (Table 2.5). The predicted size of each insert is: S39: 393 bp, S42: 393 bp, S40: 396 bp, S43: 396 bp, and the products obtained were consistent with this (Figure 6.8).

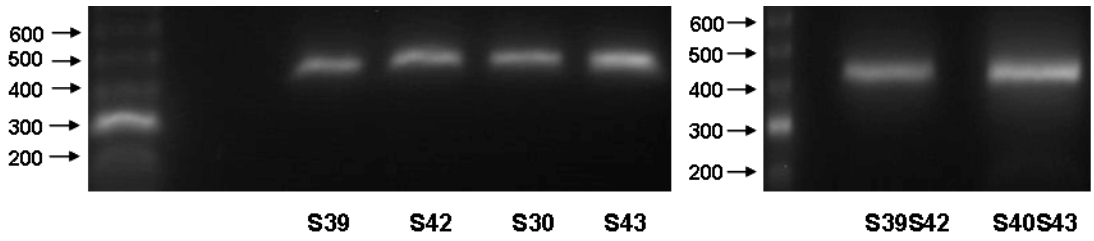


Figure 6.8 PCR amplification of the inserts. The inserts S39, S42, S39S42, S40, S43 and S40S43 were PCR amplified. The predicted size of each insert is: S39: 393 bp, S42: 393 bp, S40: 396 bp, S43: 396 bp, S39S42: 393 bp and S39S42: 396 bp.

Two rounds of PCR were performed in order to make the S39S42 insert. In the first round PCR, S39S42N (TRX3 aa 1-47) were amplified separately using the 1:100 diluted S39 PCR product as template, with the TRX3BegBmRI +TRX3^{Ser42}R primer pair. In the second round PCR, S39S42 was amplified using the 1:100 diluted S39S42N and S42C (TRX3 aa 40-119) PCR products

as templates, with the TRX3BegBmRI+TRX3EndRI primer pair (Table2.5). The S40S43 insert was made in the same way as the S39S42 insert. The predicted size of each inserts is S39S42: 393 bp, S39S42: 396 bp, and the products obtained were consistent with this (Figure 6.8).

The PCR products were digested with *EcoRI* and ligated into *EcoRI* digested RRS prey expression vector pUra. Recombinants were checked for insert size and orientation by colony PCR with a vector-specific T7 as forward primer, and an insert-specific primer e.g. TRX3EndRI as reverse primer (data not shown). Positive clones were identified and the plasmids were extracted and sequenced (data not shown). Sequencing results confirmed that all inserts were successfully cloned in frame into the *EcoRI* site in pUra, which generated the pUra-S39, pUra-S42, pUra-S39S42, pUra-S40, pUra-S43 and pUra-S40S43 constructs.

6.2.2.2 GCR1 derived baits might not interact with the S39, S42, S39S42, S40, S43 and S40S43 preys in RRS

The prey constructs pUra-S39, pUra-S42, pUra-S39S42, pUra-S40, pUra-S43, pUra-S40S43, as well as the empty prey vector pUra, were transformed separately into the bait-containing (VLCYCLFi1, CYCLF-i1, i1, i2, i1-i2 and i1-GGG-i2) and the empty bait-vector-containing (pMetRas) yeast cells. The empty bait and prey vectors were used as controls. The yeast cells were plated onto Glu-L-U+M medium and incubated at 24°C until colonies appeared. The colonies were streaked onto the Glu-L-U+M plate and incubated at 24°C until they had grown sufficiently to be replicate plated. They were then replica plated onto the selective medium Gal-L-U-M and the three control media Glu-L-U-M, Gal-L-U+4M and YPD and incubated at 36°C for 6 days. The plates were scanned every 24 hours from day 3 until day 6. Figure 6.9 shows the yeast growth on day 6.

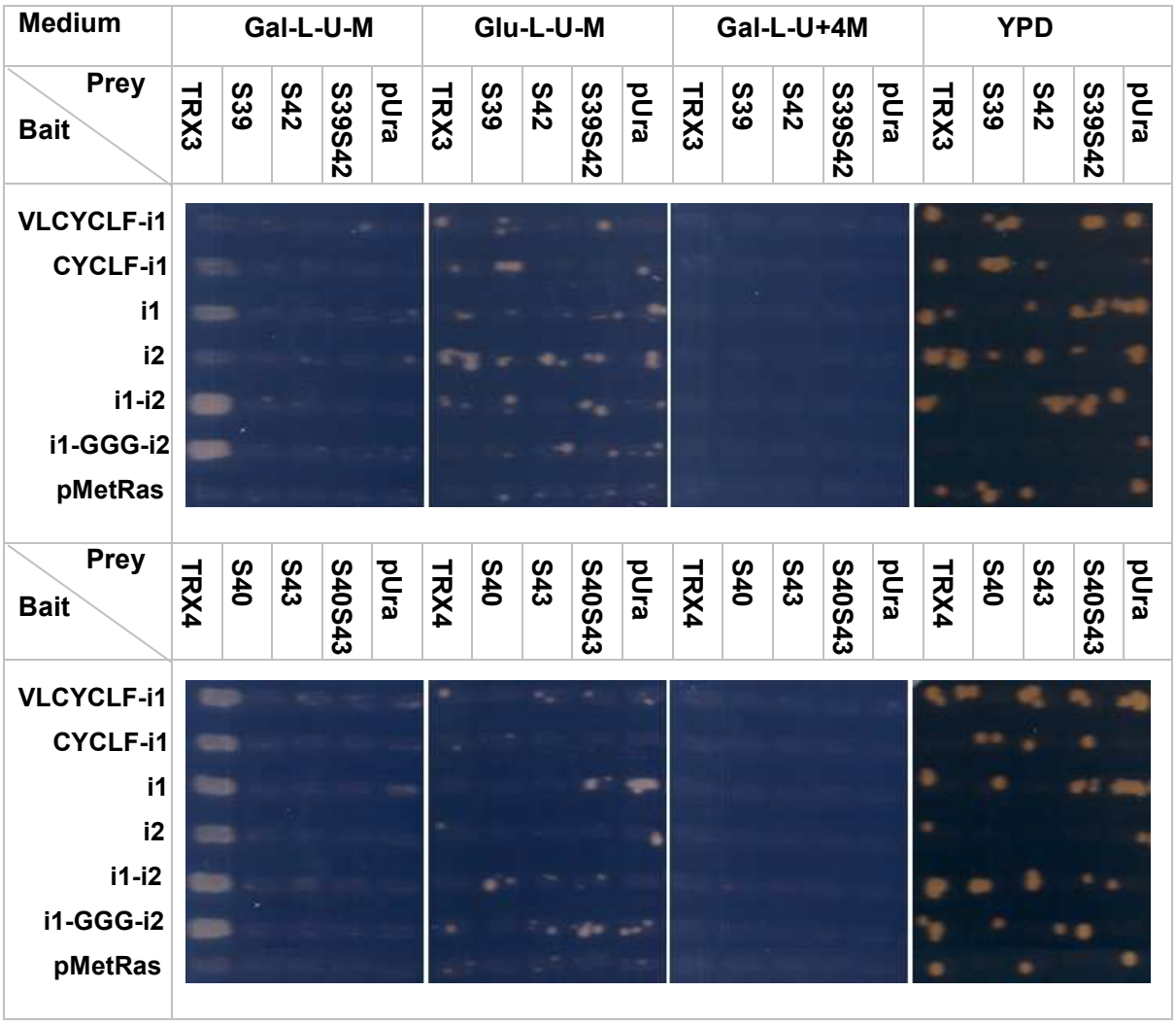


Figure 6.9 Interactions between the GCR1 derived baits and the TRX3 and TRX4 derived preys. The yeast streaks were replica plated onto the four media/plates and incubated at 36°C for 6 days. The photos show yeast growth after this 6 day period.

As shown in figure 6.9, wild type TRX3 and TRX4 interacted with all of the GCR1 derived baits. However, the TRX3 and TRX4 mutants (S39, S42, S39S42, S40, S43 and S40S43) did not show interaction with any of the baits. Prey interacting with the polypeptide encoded by the empty bait vector e.g. mRas, other than the bait, may cause yeast cell proliferation. Therefore, interactions between the prey and the empty bait vector pMetRas were used as controls (Figure 6.9). As there was no growth on the media, the bait but not the polypeptide encoded by the empty bait vector caused the interaction seen for TRX3 and TRX4. Another control that contains the empty bait vector and the

empty prey vector demonstrated that the polypeptides encoded by the empty bait and prey vectors do not interact directly to cause proliferation. The experiment was repeated one more time with four more independent transformants for each bait-prey combination, and the same result was obtained (data not shown).

6.2.3 TRX3 and TRX4 interact with all parts of GCR1 in the rRRS

This section describes the result obtained for the third objective, which was to examine whether TRX3 and TRX4 interact with other parts of GCR1 besides the i1 and i2 regions. The result was obtained by performing the rRRS screens using various parts of GCR1 as baits and TRX3 and TRX4 as preys.

6.2.3.1 Making the pMet-Nter-i2, pMet-i3-Cter and pMet-Nter-Cter rRRS bait constructs

The inserts Nter-i2, i3-Cter and Nter-Cter were PCR amplified using the Expand High Fidelity^{PLUS} PCR System (section 2.2.2.1, figure 6.10). They were amplified using *GCR1* cDNA as templates with the GCR1BegHindIII + LP2EndHindIII, LP3BegHindIII + CterEndHindIII, and GCR1BegHindIII + CterEndHindIII primer pairs respectively (Table 2.6). The predicted size of each insert is Nter-i2: 384 bp, i3-Cter: 459 bp, Nter-Cter: 1002 bp, and the PCR products obtained were consistent with this (Figure 6.10).

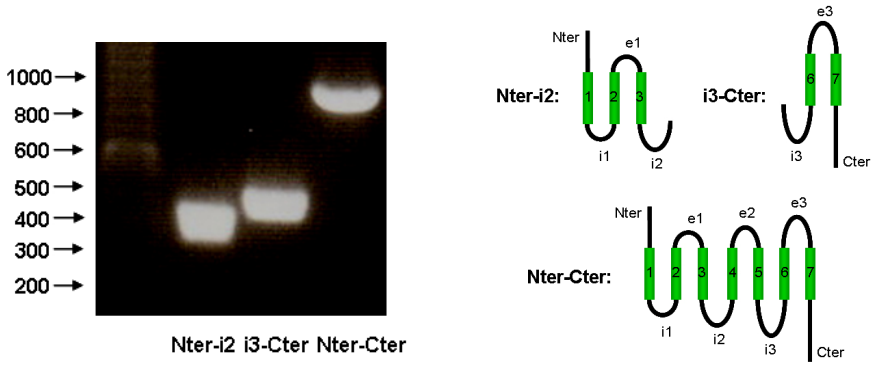


Figure 6.10 PCR amplification of the inserts. The three inserts, Nter-i2, i3-Cter, and Nter-Cter were PCR amplified. The predicted size of each insert is: Nter-i2: 384 bp, i3-Cter: 459 bp, and Nter-Cter 1002 bp.

PCR products were digested with *HindIII* (Promega) and ligated into *HindIII* digested rRRS bait expression vector pMet. Recombinants were checked for insert size and orientation by colony PCR with a vector-specific pMetF as forward primer, and an insert-specific primer e.g. LP2EndHindIII as reverse primer (data not shown). Positive clones were identified and the plasmids were extracted and sequenced (data not shown). Sequencing results confirmed that all inserts were successfully cloned in frame into the *HindIII* site in pMet, which generated the pMet-Nter-i2, pMet-i3-Cter and pMet-Nter-Cter constructs.

6.2.3.2 Making the pUraRas-TRX3 and pUraRas-TRX4 rRRS prey constructs

The TRX3 and TRX4 inserts were PCR amplified using the Expand High Fidelity^{PLUS} PCR System (section 2.2.2.1). They were amplified using *TRX3* and *TRX4* cDNA as templates with the TRX3begBmRI + TRX3EndRI and TRX4begBmRI + TRX4EndRI primer pairs respectively (Table 2.6). PCR products were digested with *EcoRI* (Promega) and ligated into the *EcoRI* digested rRRS prey expression vector pUraRas. Recombinants were checked for insert size and orientation by colony PCR with a vector-specific T7 as forward primer and an insert-specific primer e.g. TRX3EndRI as reverse primer (data not shown). Positive clones were identified and the plasmids were extracted and sequenced (data not shown). Sequencing results confirmed that all inserts were successfully cloned in frame to the C-terminus of Ras in pUraRas (*EcoRI* site), which generated the pUraRas-TRX3 and pUraRas-TRX4 constructs.

6.2.3.3 Autoactivation test of the TRX3 and TRX4 rRRS preys

The pUraRas-TRX3 and pUraRas-TRX4 constructs were transformed into the temperature sensitive yeast strain *cdc25-2*. Four yeast colonies for each of the preys (TRX3 and TRX4) were randomly selected for the temperature sensitivity test. They were streaked onto a Glu-U+M plate and incubated at 24°C for 3

days. They were then replica plated onto two Glu-U+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as a control. There was virtually no yeast growth at 36°C, but all yeast streaks grew at 24°C (Figure 6.11). Therefore these prey-containing yeast cells were temperature sensitive and were used in the following experiments.

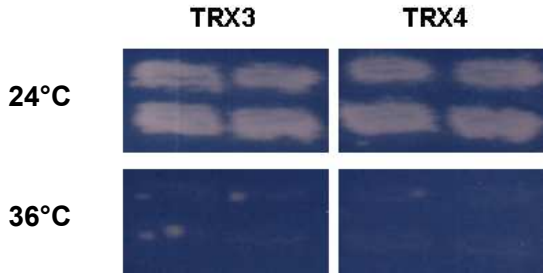


Figure 6.11 Prey temperature sensitivity test. The prey-containing (TRX3 or TRX4) yeast cells were streaked onto Glu-U+M plate and incubated at 24°C for 3 days. They were then replica plated onto two Glu-U+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as a control. Photos show yeast growth after this 3 day period.

The rRRS prey vector pUraRas contains mRas, which might be brought to the cell membrane when it is fused to the prey, resulting in autoactivation of the system. It is vital to work only with preys that do not autoactivate the rRRS. To check for autoactivation, yeast cells expressing prey with no bait were replica plated onto two Gal-U+M plates, one was grown at 36°C for 5 days, and the other at 24°C for 5 days as a control. There was virtually no yeast growth at 36°C, but all yeast streaks grew at 24°C (Figure 6.12). Therefore the TRX3 and TRX4 preys did not autoactivate, and were used in the following rRRS screens.

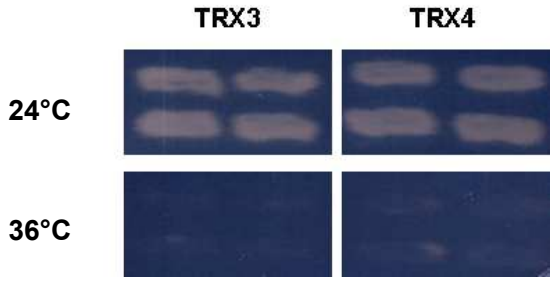
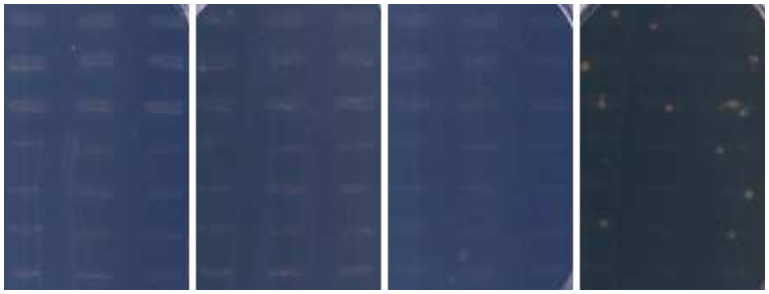
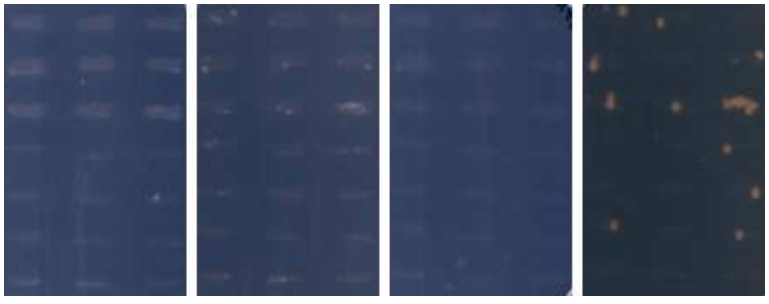


Figure 6.12 Prey autoactivation test. To check whether the TRX3 and TRX4 preys autoactivate the rRRS or not, the yeast cells were replica plated onto two Gal-U+M plates,

one was grown at 36°C for 5 days, and the other at 24°C for 5 days as a control. Photos show yeast growth after this 5 day period.

6.2.3.4 TRX3 and TRX4 interact with all parts of GCR1 in the rRRS

The pMet-Nter-i2, pMet-i3-Cter and pMet-Nter-Cter rRRS bait constructs were co-transformed separately with pUraRas-TRX3, pUraRas-TRX4 and pUraRas into the temperature sensitive yeast strain *cdc25-2*. The yeast cells were plated onto Glu-L-U+M plates and incubated at 24°C until colonies appeared. Three randomly selected colonies for each of the bait-prey combination were streaked onto the Glu-L-U+M plates and incubated at 24°C until they had grown sufficiently to be replicate plated. They were then replica plated onto the selective medium Gal-L-U-M and the three control media Glu-L-U-M, Gal-L-U+4M and YPD and incubated at 36°C for 6 days. The plates were scanned every 24 hours from day 3 until day 6 (Figure 6.13 and 6.14).

Expression / Bait + Prey Media	Bait + Prey	Bait	Prey	None
	Gal-L-U-M	Glu-L-U-M	Gal-L-U+4M	YPD
	Day 3			
Nter-i2 + TRX3 i3-Cter + TRX3 Nter-Cter + TRX3 Nter-i2 + pUraRas i3-Cter + pUraRas Nter-Cter + pUraRas pMet + pUraRas				
	Day 4			
Nter-i2 + TRX3 i3-Cter + TRX3 Nter-Cter + TRX3 Nter-i2 + pUraRas i3-Cter + pUraRas Nter-Cter + pUraRas pMet + pUraRas				

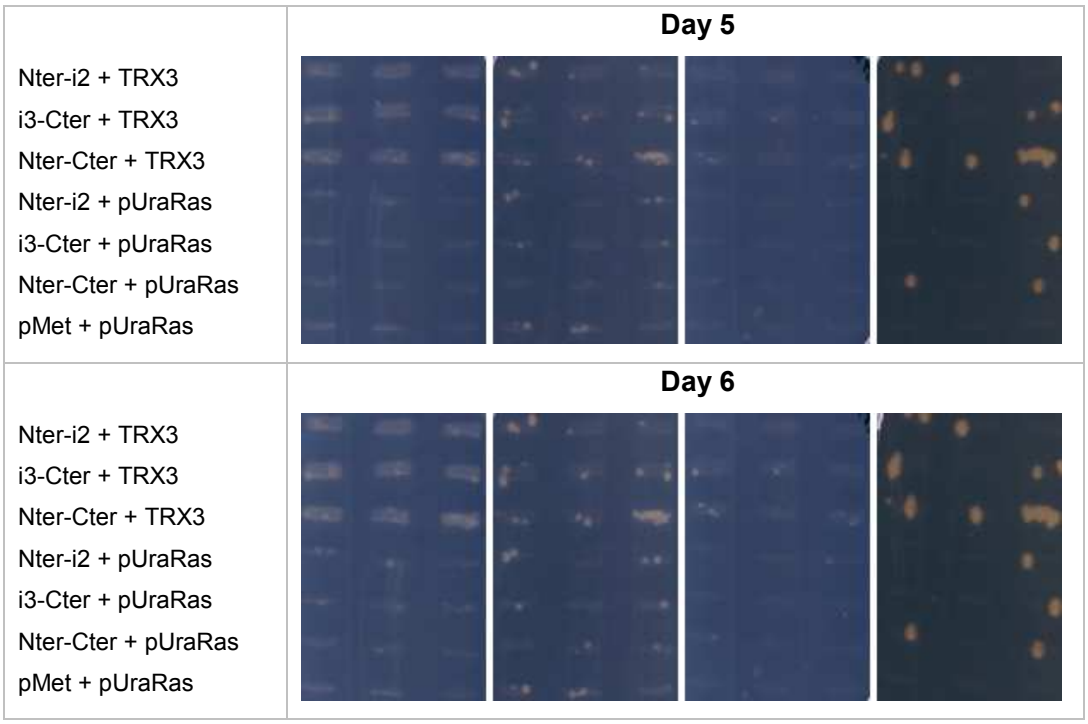


Figure 6.13 Interactions between the GCR1 derived baits and TRX3 in the rRRS. The yeast streaks were replica plated onto the four media/plates and incubated at 36°C for 6 days. The plates were scanned every 24 hours from day 3 to day 6.

Over the course of the experiment, definite growth can be seen on Gal-L-U-M media for all interaction between GCR1 derived baits (Nter-i2, i3-Cter and Nter-Cter) and TRX3 (Figure 6.13, row 1-3). It should be noted that there were three transformants for each bait-prey interaction. The interactions seen for the three transformants were identical, the result was hence reproducible. Bait interacting with the polypeptide encoded by the empty prey vector e.g. mRas, other than the prey, may cause yeast cell proliferation. Therefore, interactions between the GCR1 derived baits and the empty rRRS prey vector pUraRas (Figure 6.13, row 4-6) were used as controls. As there was no growth on the media, the prey but not the polypeptide encoded by the empty prey vector caused the interaction seen for TRX3. Another control that contains the empty bait vector and the empty prey vector (Figure 6.13, row 7) demonstrated that the polypeptides encoded by the vectors do not interact directly to cause proliferation. Altogether, above result indicates that TRX3 interacts with all parts

of GCR1, confirming that the GCR1-TRX3 interaction first seen in the RRS is likely to be true.

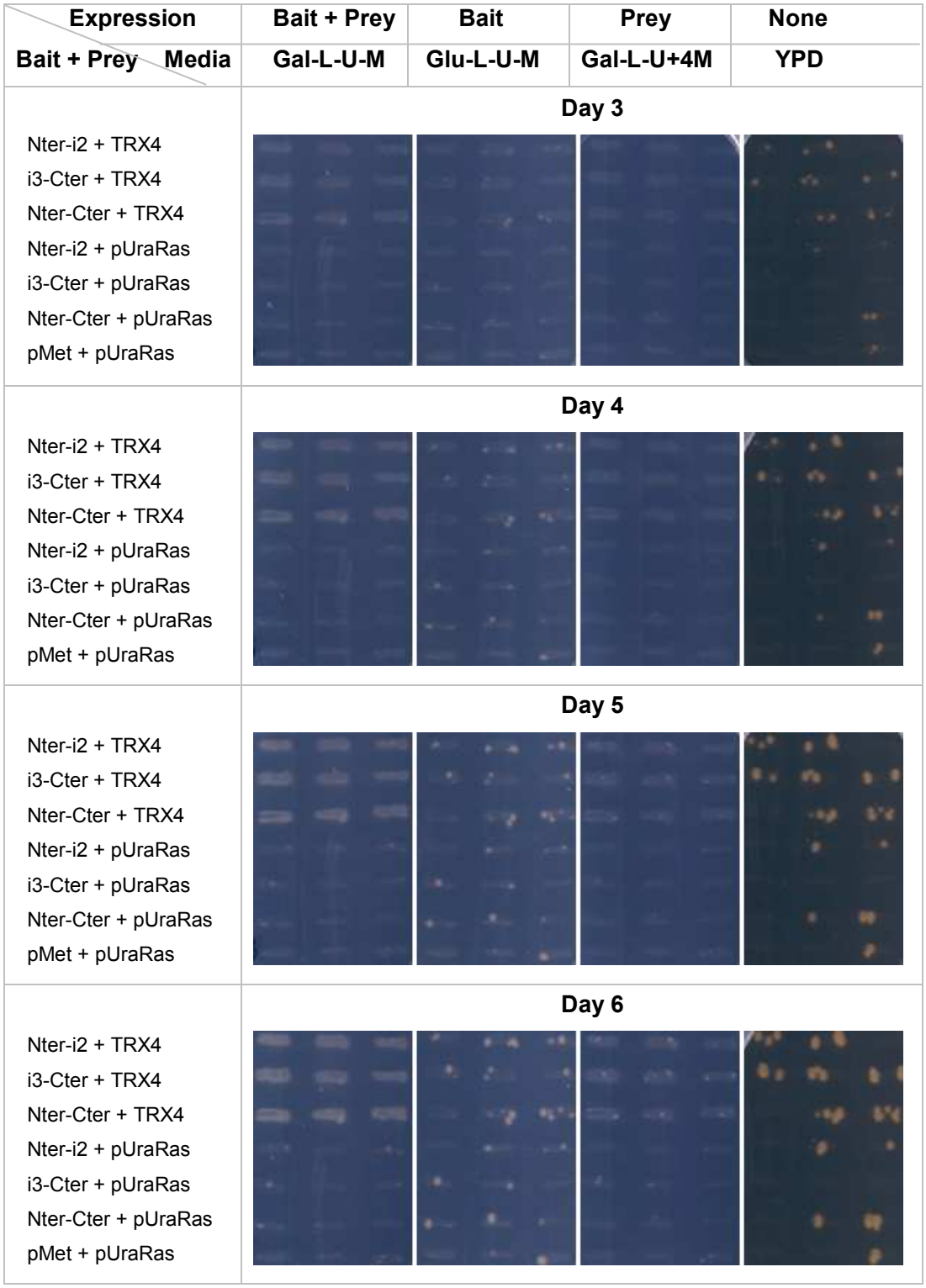


Figure 6.14 Interactions between GCR1 derived baits and TRX4 in the rRRS. The

yeast streaks were replica plated onto the four media/plates and incubated at 36°C for 6 days. The plates were scanned every 24 hours from day 3 to day 6.

Over the course of the experiment, definite growth can be seen on Gal-L-U-M media for the interactions between the Nter-i2, i3-Cter and Nter-Cter baits and TRX4 (Figure 6.14, row 1-3). The same as for TRX3, the interactions seen for the three independent transformants of TRX4 were identical, indicating that the result was reproducible. Also, there was no growth over the course of the experiment for the controls (Figure 6.14, row 4-7), therefore the prey but not the polypeptide encoded by the empty prey vector caused the interaction seen for TRX4. The result indicates that TRX4 interacts with all parts of GCR1, confirming that the GCR1-TRX4 interaction first seen in the RRS is likely to be true. It should also be noted that the growth for TRX4 appeared to occur slightly earlier than TRX3, as seen in the RRS (Figure 6.7). In addition, a low level of growth is visible on the Gal-L-U+M media (Figure 6.14), especially for day 5 and day 6.

6.2.4 Reconfirming the GCR1-TRX3 and GCR1-TRX4 interactions using the pull-down assays

This section describes the result obtained for the last objective, which was to verify the interactions seen in the RRS using the pull-down assays, with the 6xHis-tagged i1-GGG-i2 as bait and Myc-tagged TRX3, TRX4, S42 and S43 as preys.

6.2.4.1 Cloning, expression and purification of the i1-GGG-i2-6xHis bait

The pMetRas-i1-GGG-i2 plasmid was digested with *NcoI* (NEB, section 2.2.2.2) in order to release the i1-GGG-i2 fragment. The fragment of the expected size for i1-GGG-i2 (111 bp, figure 6.15) was gel purified. It was ligated into the *NcoI* digested *E.coli* expression vector pET22b(+) (Table 2.2, figure 2.2). Recombinants were checked for insert size and orientation by colony PCR with

a vector-specific T7 as forward primer and an insert-specific LP2EndEcoRI as reverse primer (data not shown). Positive clones were identified and the plasmids were extracted and sequenced (data now shown). Sequencing results confirmed that i1-GGG-i2 was successfully cloned in frame to the N-terminus of the 6xHis sequence in pET22b(+) (*NcoI* site), which generated the pET22b(+)-i1-GGG-i2 construct.

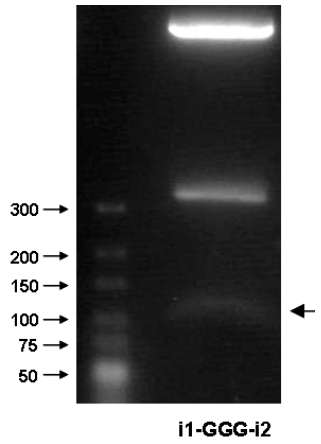


Figure 6.15 Preparation of the i1-GGG-i2 insert. The pMetRas-i1-GGG-i2 plasmid was digested by *NcoI* to release the i1-GGG-i2 fragment. The fragment of the expected size for i1-GGG-i2 (111 bp) was gel purified.

The pET22b(+)-i1-GGG-i2 construct was transformed into the *E.coli* expression strain BL21 (DE3). The expression of the recombinant protein was induced (section 2.2.13.1) by 0.5 mM IPTG at 37°C for 3, 4, 5 and 6 hours respectively and assessed by analysis of the total cell fraction on a SDS-PAGE followed by Coomassie blue staining (section 2.2.13.2). The expression level of the recombinant protein i1-GGG-i2-6xHis was almost identical when induced for 3, 4, 5 and 6 hours (Figure 6.16A), therefore induction of the recombinant protein expression was performed for 3 hours in subsequent experiments. The solubility of the recombinant protein was checked as described in section 2.2.13.4, and it was found to be expressed in *E. coli* as soluble protein (Figure 6.16B).

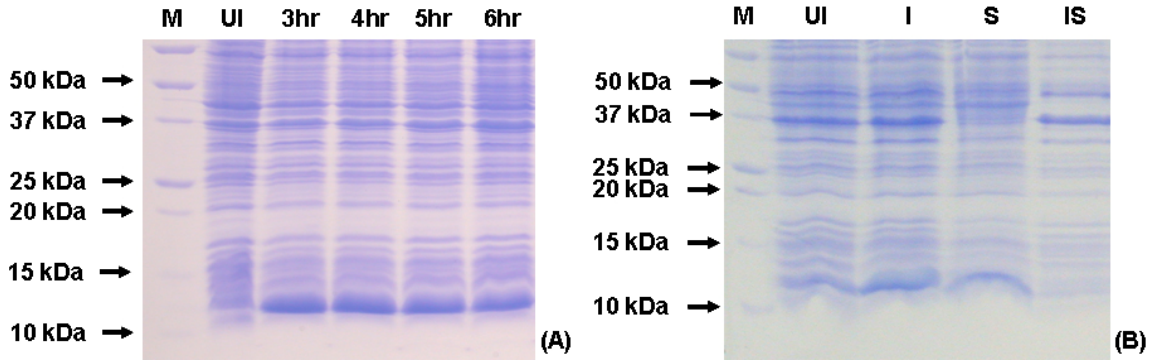


Figure 6.16 Expression (A) and solubility (B) check of the recombinant protein i1-GGG-i2-6xHis. The expression of the recombinant protein was induced by 0.5 mM IPTG at 37°C for 3 hours (3hrs), 4 hours (4hrs), 5 hours (5hrs) and 6 hours (6hrs) respectively. Protein solubility check was performed as described in section 2.2.13.4. UI: uninduced total cell fraction; I: induced total cell fraction; S: soluble fraction; IS: insoluble fraction. The predicted size of the recombinant protein is 9 kDa.

The recombinant protein was purified using His-affinity beads as described in section 2.2.14. The uninduced total cell fraction, input (cell lysate containing expressed i1-GGG-i2-6xHis), supernatant of the binding step, supernatant of the first two rounds of washing step, together with the elution fraction were assessed by SDS-PAGE. As shown in figure 6.17, a polypeptide of the expected size for i1-GGG-i2-6xHis were present only in the input and elution fraction, which indicates that i1-GGG-i2-6xHis was be able to bind to the His-affinity beads and eluted under suitable conditions (section 2.2.15).

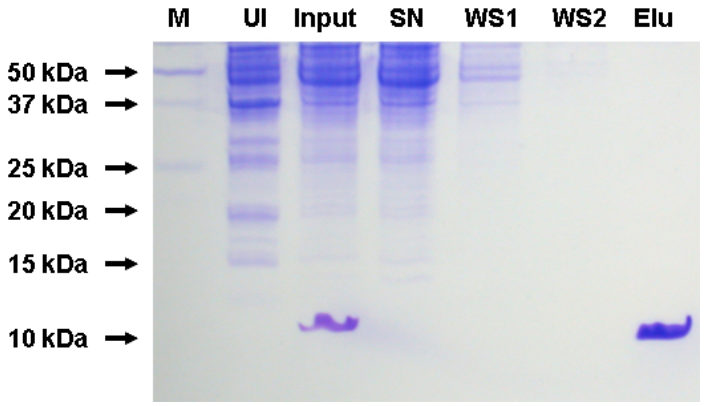


Figure 6.17 Purification of i1-GGG-i2-6xHis using His-affinity beads. UI: uninduced total cell fraction; Input: cell lysate containing expressed i1-GGG-i2-6xHis; SN: supernatant of the binding step; WS1: supernatant of the first washing step; WS2: supernatant of the second washing step; Elu: elution fraction. 1/10 equivalent of the input

sample volume was used for the UI, Input, SN, WS1 and WS2 fractions, whereas 1/2 equivalent of the input sample volume was used for the Elu fraction in this SDS-PAGE.

6.2.4.2 Cloning and expression of the Myc-TRX3, Myc-TRX4, Myc-S42 and Myc-S43 preys

The inserts (TRX3, TRX4, S42 and S43) were amplified by PCR using the Expand High Fidelity^{PLUS} PCR System (section 2.2.2.1, figure 6.18). TRX3 and TRX4 were amplified using *TRX3* and *TRX4* cDNA as templates, with the TRX3BegEcoRI + TRX3EndRI and TRX4BegEcoRI + TRX4EndRI primer pairs respectively (Table2.7). The S42 and S43 inserts were amplified using pUra-S42 and pUra-S43 as templates, with the TRX3BegEcoRI + TRX3EndRI and TRX4BegEcoRI + TRX4EndRI primer pairs respectively (Table 2.7). The predicted size of each insert is TRX3: 386 bp, TRX4: 389 bp, S42: 386 bp, S43: 389 bp, and the PCR products obtained were consistent with this (Figure 6.18).

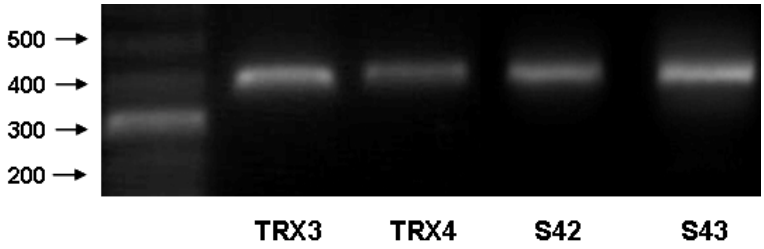


Figure 6.18 PCR amplification of the insert. The TRX3, TRX4, S42, S43 inserts were PCR amplified. The predicted size of each insert is: TRX3: 386 bp, TRX4: 389 bp, S42: 386 bp, S43: 389 bp.

The PCR products were digested with *EcoRI* and ligated into the *EcoRI* digested *E.coli* expression vector pET22b(+)myc (Table 2.2, figure 2.3). Recombinants were checked for insert size and orientation by colony PCR with a vector-specific T7 as forward primer, and an insert-specific primer e.g. TRX3EndRI as reverse primer (data not shown). Positive clones were identified and the plasmids were extracted and sequenced (data now shown). Sequencing results confirmed that all inserts were successfully cloned in frame to the C-terminus of Myc in pET22b(+)myc (*EcoRI* site), which generated the

pET22b(+)myc-TRX3 (Myc-TRX3), pET22b(+)myc-TRX4 (Myc-TRX4), pET22b(+)myc-S42 (Myc-S42) and pET22b(+)myc-S43 (Myc-S43) constructs.

All of the above constructs were transformed into the *E.coli* expression strain BL21 (DE3). The expression of the recombinant proteins was induced (section 2.2.13.1) by 0.5mM IPTG at 25°C for 3 hours and assessed by analysis of the total cell fraction on a SDS-PAGE followed by Coomassie blue staining (section 2.2.13.2, figure 6.19). The solubility of the recombinant protein was checked as described in section 2.2.13.4. All of the four recombinant proteins were found to be highly expressed in *E. coli* as soluble proteins (Figure 6.19).

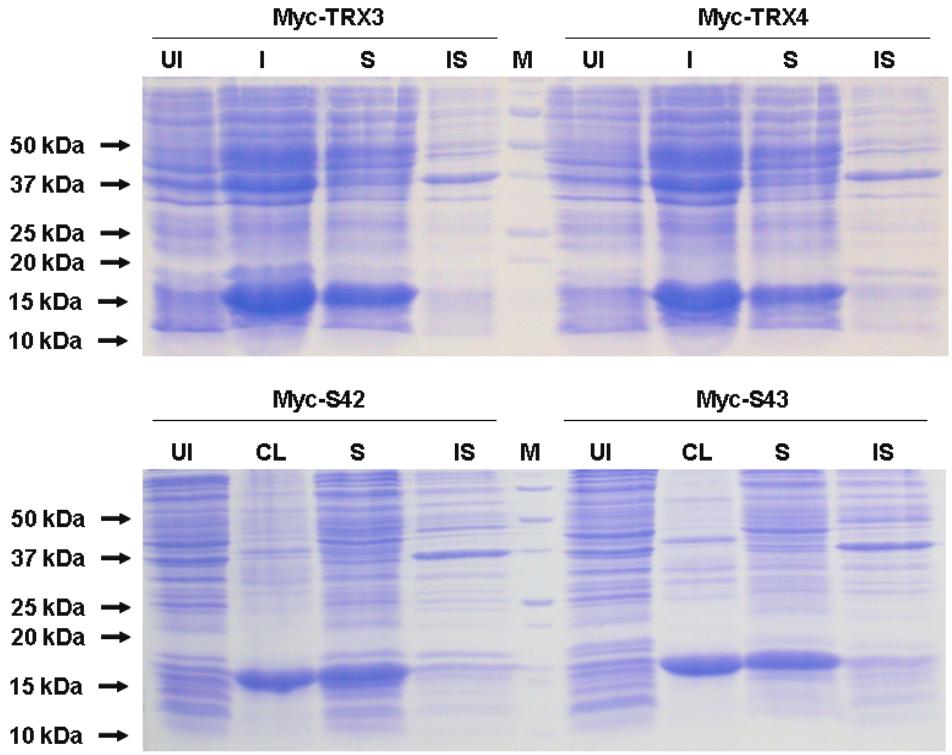


Figure 6.19 Expression and solubility check of the recombinant proteins Myc-TRX3, Myc-TRX4, Myc-S42 and Myc-S43. The expression of the recombinant proteins was induced by 0.5mM IPTG at 25°C for 3 hours. Protein solubility check was performed as described in section 2.2.13.4. UI: uninduced total cell fraction; I: induced total cell fraction; CL: cell lysate; S: soluble fraction, IS: insoluble fraction. The predicted size of each recombinant protein is: Myc-TRX3: 16.94 kDa, Myc-TRX4: 16.90 kDa, Myc-S42: 16.94 Dka, Myc-S43: 16.94 kDa

6.2.4.3 i1-GGG-i2-6xHis interacts with Myc-TRX4 and Myc-S43 in the pull-down assays.

Pull-down assays were performed (section 2.2.15) using the 6xHis-tagged i1-GGG-i2 as bait and Myc-tagged TRX3, TRX4, S42 and S43 as preys. For each pull-down assay, one experiment group and two control groups were included. The experiment group contains the i1-GGG-i2-6xHis bait and the prey e.g.Myc-TRX3. The first control group contains the polypeptide (6xHis) encoded by the empty bait vector, and the prey. The second control group contains the prey alone. The control groups were used to test whether the prey would non-specifically bind to the polypeptide encoded by the empty bait vector or to the His-affinity beads. The input and elution fractions of each group were analysed by SDS-PAGE and Coomassie staining to detect the bait (Figure 6.20-21A), and by Western blot with anti-Myc antibody to detect the prey (Figure 6.20-21B).

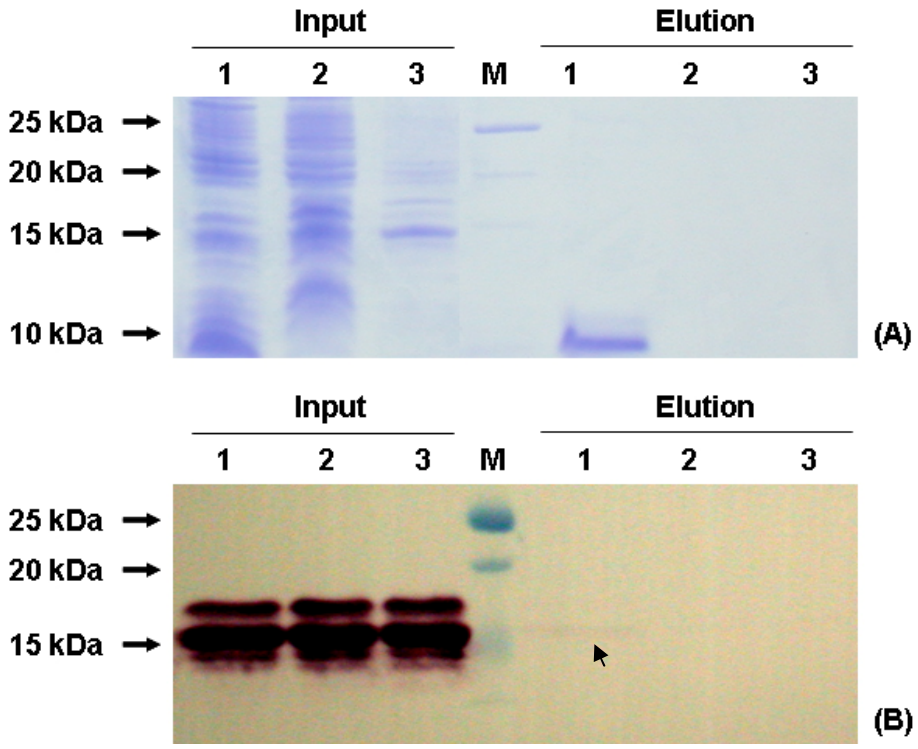


Figure 6.20 Pull-down assay for i1-GGG-i2-6xHis with Myc-TRX4. (A) Coomassie staining detecting bait. (B) Western blot with the anti-Myc antibody detecting prey. Lane 1: i1-GGG-i2-6xHis with Myc-TRX4, Lane 2: 6xHis with Myc-TRX4; Lane 3: Myc-TRX4 only.

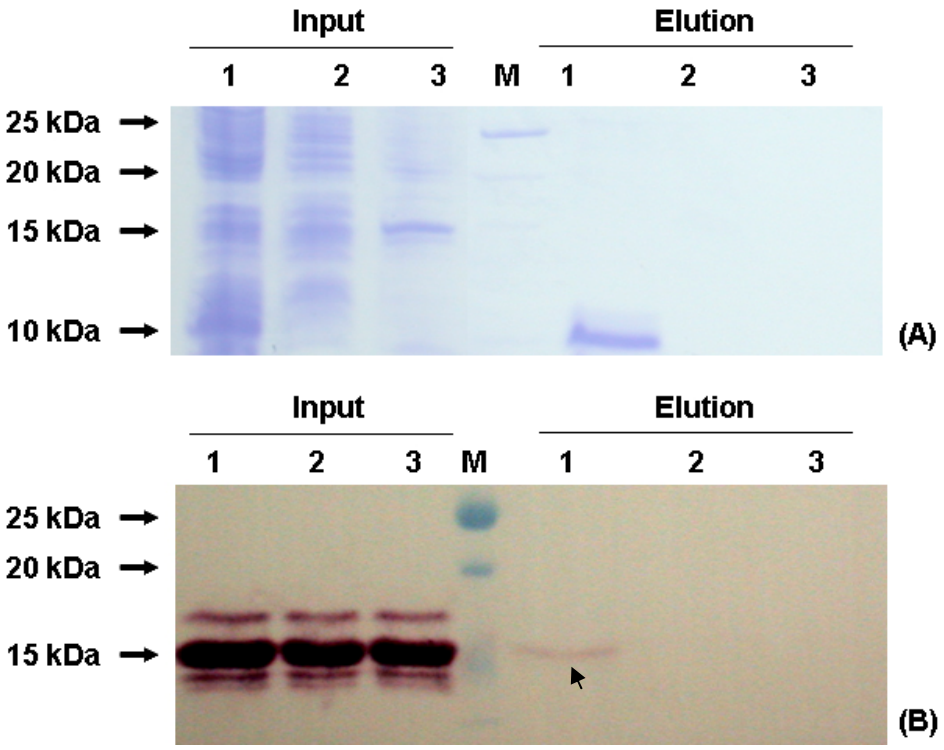


Figure 6.21 Pull-down assay for i1-GGG-i2-6xHis with Myc-S43. (A) Coomassie staining detecting bait. (B) Western blot with anti-Myc antibody detecting prey. Lane 1: i1-GGG-i2-6xHis with Myc-S43, Lane 2: 6xHis with Myc-S43; Lane 3: Myc-S43 only.

As shown in figure 6.20A, there was a polypeptide of the predicted size for i1-GGG-i2-6xHis in lane 1, which indicates that i1-GGG-i2-6xHis was eluted from the His-affinity beads. As shown in figure 6.20B, there was a faint band of the predicted size for Myc-TRX4 present in lane 1, which indicates that Myc-TRX4 was co-eluted with i1-GGG-i2-6xHis. There was no band of the expected size for Myc-TRX4 presented in lane 2 and 3 (Figure 6.20B), indicating that there was no non-specific binding of the prey to the polypeptide encoded by the empty bait vector (lane 2) or to the His-affinity beads (lane 3). Taken together, TRX4 interacts weakly but specifically with i1-GGG-i2 in the pull-down assay. The same result was obtained for S43 (Figure 6.21), indicating that it also interacts specifically with i1-GGG-i2 in the pull-down. Comparing the band in lane 1 of the elution fraction in figure 6.20B with the corresponding band in figure 6.21B, the latter is more intense than former. The experiment was repeated one more time, and the same result was obtained. Although the experiment was not quantitative, it may suggest that compared with TRX4,

there was more Myc-S43 co-eluted with i1-GGG-i2. Thus the interaction for i1-GGG-i2 with S43 might be stronger than with TRX4. No band of the expected size for Myc-TRX3 and Myc-S42 could be detected in the elution fraction in the experimental and control groups (data not shown).

6.3 Discussion

6.3.1 The GCR1-TRX3 and GCR1-TRX4 interactions in the RRS

Two thioredoxins, TRX3 and TRX4, were identified in the initial library screening as GCR1 interactors (chapter 4). They were both reconfirmed to interact with the i1 and i2 regions of GCR1 in the RRS (chapter 5). Thioredoxins are known to interact with cysteine or methionine residues, but neither of the two loops contains cysteine or methionine residues. Literature searches revealed that there are thioredoxin substrates which do not have cysteines. For instance, 20 of the 80 proteins identified to interact with thioredoxin by Kumar *et al* (2004) do so independent of mixed-disulfide formation because they do not contain cysteines. Therefore the identification of i1 and i2 as thioredoxin substrates is reasonable. Interestingly, there are two cysteines in the motif VLCYCLF in the TM1 region near i1 (Figure 3.1). As the boundaries of the intracellular domains may move up and down in relation to the membrane when the protein is in different conformations, the two cysteines could be exposed to the cytoplasm, becoming a potential binding site for thioredoxins. If thioredoxins interact with the motif containing the cysteines, the strength of the interaction for TRX3 and TRX4 with the extended i1 baits containing these cysteines (VLCYCLF-i1 and CYCLF-i1) is expected to be greater than with the i1 bait. Therefore we tested the interaction for TRX3 and TRX4 with the VLCYCLF-i1 and CYCLF-i1 baits. The i1, i2, i1-i2 and i1-GGG-i2 baits as well as the TRX2 prey were included as controls in this experiment.

The stronger the interaction between the bait and prey, the more likely it is that the mRas will be membrane localised quickly and retained long enough to

induce cell proliferation. It is therefore assumed that a high level of yeast growth implicates a strong degree of interaction between the bait and prey. As shown in figure 6.7, yeast growth level was lower for VLCYCLF-i1 compared with CYCLF-i1 and i1, suggesting a weaker interaction. However, the Western blot analysis shows that the baits are not expressed at the same level (Figure 6.4) so the differences in yeast growth seen may not be due entirely to the strength of the protein-protein interaction. In fact, VLCYCLF-i1 was expressed at a much lower level than CYCLF-i1 and i1 (Figure 6.4), therefore the lower level of growth for VLCYCLF-i1 may due to lower protein concentration rather than weaker interaction. It is possibly that the interactions for TRX3 and TRX4 with VLCYCLF-i1 are even higher than with CYCLF-i1 and i1. As for CYCLF-i1, it was expressed at a much lower level than i1, but only displayed a slightly lower yeast growth level than i1 (Figure 6.7), indicating that its interaction with TRX3 and TRX4 may be stronger than i1. Another possibility that should be considered is that the level of interaction may be limited by the amount of prey present in the yeast cells, which could lead to saturation of the growth response. Given that the level of yeast growth varied for different baits, it is more likely that there was sufficient prey, and that the level of yeast growth was a reflection of interaction strength for individual bait. Altogether, the result leads us to tentatively suggest that TRX3 and TRX4 could bind to the cysteines in TM1 where these are included in the VLCYCLF-i1 and CYCLF-i1 baits and reduce one or both of these cysteines. The result of interactions with extended i1 regions fused to the i2 region (VLCYCLF-i1-i2 and CYCLF-i1-i2, table 2.5) may have provided more information, but unfortunately both showed autoactivity (data not shown).

The observation that the i2 bait fusion protein was expressed at a much higher level than i1 (Figure 6.4), and that i2 had a much lower level of yeast growth than i1 (Figure 6.7) indicates that TRX3 and TRX4 bind preferentially to i1 than i2 during a GCR1 interaction. Furthermore, the interaction may have been enhanced when both i1 and i2 were present, since the level of yeast growth for i1-i2 and i1-GGG-i2 were higher than for i1 and i2, and their interactions

occurred earlier than i1 and i2 (Figure 6.7). However, it is also possible that the relatively slower yeast growth for i1 compared with i1-i2 and i1-GGG-i2 was due to its lower expression level than i1-i2 and i1-GGG-i2 (Figure 6.4), as its growth did come up to the same level as i1-i2 and i1-GGG-i2 after day 4 (Figure 6.7). Again, insufficient prey limiting the interaction was unlikely to have happened, as the level of growth varied, especially on the day 3 and day 4 (Figure 6.7).

Overall, TRX3 and TRX4 displayed similar binding profiles to the GCR1 derived baits. This may reflect the fact that both TRX3 and TRX4 belong to subgroup 1 of *Arabidopsis* thioredoxin *h* family, so could possibly share the same interactions. Besides, the growth for the TRX4 appeared to occur slightly earlier than TRX3. Although the experiment is not quantitative, it might indicate that TRX4 has a stronger GCR1 interaction than TRX3. Interestingly, TRX2 did not interact with any of the baits except i1-i2. This was in accordance with where it came out from the initial library screening (i1-i2), and was identical to the verification result shown in figure 5.7. TRX2 hence had served as a good negative control for TRX3 and TRX4 in this experiment. The possible reason that TRX3 and TRX4, but not TRX2, interact with GCR1 is addressed in section 6.3.4.

One of the drawbacks of using the yeast two-hybrid system to analyse protein-protein interaction is that the experiment is not quantitative, because we can not precisely measure and control the level of bait and prey proteins present in the yeast cell. However, this drawback can be overcome by employing other methods, such as surface plasmon resonance (SPR), which allows the detection of protein-protein in real time by measuring the equilibrium binding constants as well as the association and dissociation rates (Slepak, 2000; Seitz *et al*., 2006). SPR has been successfully used to investigate GPCR signalling, such as GPCR-agonist interactions (Alves *et al*., 2005; Harding *et al*., 2006). It could in the future be used to measure the strength of the interactions for VLCYCLF-i1, CYCLF-i1 and i1 with TRX3 and TRX4, so that the differences

among these interactions can be compared and the contribution of the cysteine residues to the interactions can be analysed.

6.3.2 The GCR1-TRX3 and GCR1-TRX4 interactions in the rRRS

The rRRS screens were performed in order to examine if TRX3 and TRX4 interact with other parts of GCR1 (section 6.2.3.3). Both TRX3 (Figure 6.13) and TRX4 (Figure 6.14) interacted with each half of the GCR1 protein (Nter-i2 and i3-Cter), as well as the entire protein (Nter-Cter). This indicates that GCR1 may have multiple binding sites for thioredoxins instead of one area, as the result from the RRS experiment implicated. While the i3 and Cter regions do not contain any cysteines they do contain a number of methionines, which could be potential binding sites for thioredoxins. If the Nter-i2 region is incorporated into the yeast cell membrane in the same conformation predicted for the intact protein it would contain no methionine residues exposed to the cytoplasm and the cysteine residues in TM1 may or may not be exposed to the cytoplasm. This might suggest that interactions for the i3-Cter and Nter-Cter baits could be stronger than with the Nter-i2 bait. However, the level of yeast growth was almost identical for the three baits, indicating that the interaction levels might be the same. This may suggest that the thioredoxins is not targeting the GCR1 protein with its reductive capabilities and could instead be binding to provide other roles or functions to GCR1 as covered in section 6.3.5. Alternatively, it may due to the differences in bait or prey concentrations. Because both the bait and prey vectors (for the rRRS) do not contain any epitope tag, we could not check the expression levels of baits and preys. Nevertheless, the full length GCR1, which we assume can be localised to the cell membrane in its native conformation, has been shown to interact with TRX3 and TRX4 (Figure 6.13 and 6.14). Therefore the rRRS reconfirmation is a valuable support to the RRS result.

A low level of yeast growth (Figure 6.13 and 6.14) is visible on the Gal-L-U+4M media on which only prey is expected to be expressed. This is especially true

for days 5 and 6. This could be expected as methionine can not completely suppress bait expression (section 3.3.2), and the low level of expressed bait protein binding to the prey would lead to cell proliferation.

6.3.3 The GCR1-TRX3 and GCR1-TRX4 interactions *in vitro*

Although GCR1 has been confirmed to interact with TRX3 and TRX4 in both the RRS and the rRRS, it is wise to reconfirm the interactions using an independent method. Typically, the first independent assay to be used to reconfirm an interaction is a pull-down assay. Ideally, one would perform a pull-down using a full length protein. The difficulty of getting the full length GCR1 expressed as a soluble protein in *E.coli* has prohibited using the full length protein in the pull-downs. An alternative would be to use i1-GGG-i2, which was used as a successful bait in the RRS, to substitute the full length GCR1. The 6xHis tagged i1-GGG-i2 can be expressed as a soluble protein in *E.coli* (Figure 6.16) and purified using the His-affinity beads (Figure 6.17), so it was used as the bait in the pull-down assays.

The premise of the pull-down was that if the prey, e.g.Myc-TRX4, interacts with the bait i1-GGG-i2-6xHis, it would be co-purified with the bait by the His-affinity beads. The presence of the prey in the elution fraction indicates its interaction with the bait. However, it might also be caused by the prey binding non-specifically to the polypeptide encoded by the empty bait vector or to the His-affinity beads. As shown in figure 6.20 and 6.21, both TRX4 and S43 could be co-eluted with i1-GGG-i2, and they were not present in the elution fractions of the two control groups, therefore the binding of TRX4 (Figure 6.20) and S43 (Figure 6.21) to i1-GGG-i2 was specific. Furthermore, there was more S43 co-eluted with i1-GGG-i2 compared with TRX4 (section 6.2.4.3), thus the interaction for i1-GGG-i2 with S43 might be stronger than with TRX4. This may be due to the possible conformational changed caused by the cysteine to serine mutation in the active site, or due to the possible high affinity of the serine residue to i1-GGG-i2. However, no band for the TRX3 and S42 prey was

present in the elution fractions of the experimental and control groups (section 6.2.4.3). Given that the GCR1-TRX3 interaction was weaker than the GCR1-TRX4 interaction as demonstrated by the RRS and rRRS, and that the amount of TRX4 or S43 co-eluted with i1-GGG-i2 in the pull-down was very low, it is possible that TRX3 or S42 could have been co-eluted with i1-GGG-i2 but the amount of protein present in the elution fraction was too low to be detected by the Western blot. Alternatively, they did not interact with i1-GGG-i2 in the pull-down assay and were thus not co-eluted. It is worth to mention that in order to track whether TRX3 and TRX4 interact with other intracellular regions of GCR1 *in vitro*, we also made the 6xHis-CYCLF-i1, 6xHis-i1, 6xHis-i2, 6xHis-i3, 6xHis-Cter and 6xHis-Nter-i2 constructs (Table 2.7), but we did not have enough time to perform pull-downs using these baits. Given more time, performing pull-downs with these baits would be appropriate and the result would provide more information about their interactions.

In summary, GCR1 interacts with TRX3 in the RRS and rRRS, while GCR1 interacts with TRX4 in the RRS, rRRS and in the pull-down assay. Two obvious questions that come up along with these observations are: how do TRX3 and TRX4 interact with GCR1 and why do they interact. Base on the results we have obtained so far, together with a comprehensive literature study, we raised several hypotheses that are addressed in the next two sections.

6.3.4 The possible interaction mechanisms for GCR1 with TRX3 and TRX4

The interaction between thioredoxin and its target is suggested to involve the active site WCXPC, which reduce the disulfide bond formed by two cysteine residues or the methionine sulfoxide on the target protein. This is achieved through a two-step process (Holmgren, 1995; Verdoucq *et al*., 1999). In the first step, the first cysteine in the active site of thioredoxin attacks and reduces the disulfide bridge of the target protein, establishing a disulfide bridge with one cysteine of the target that forms a thioredoxin-target intermediate (Verdoucq *et*

al., 1999). In the second step, the second cysteine of the thioredoxin attacks the intermediate disulfide bridge and releases the reduced target protein. Substitution of one or both of the active site cysteines by its structural analog, e.g. serine, could abolish the interaction between thioredoxin and its target protein. For example, p40phox (a component of phagocyte oxidase) was shown to interact with wild type human thioredoxin; in contrast, no interaction was observed with double mutant C32S/C35S, which lacks reducing activity (Nishiyama *et al*, 1999). However, accumulating evidence reveals that substitution of the second cysteine in the active site stabilises the thioredoxin-target intermediate. For instance, unlike wild type thioredoxin, the C35S mutant constitutively binds to ASK-1 (Liu and Min, 2002). The Cys-46 mutant of thioredoxin *f* and Cys-155 mutant of spinach chloroplast frustose 1,6-biophosphatase mutant could form a stable mixed disulfide intermediate (Balmer and Schurmann, 2001).

As stated in section 6.3.1, TRX3 and TRX4 may bind to and reduce one or both of the cysteines in TM1 where these are included in the VLCYCLF-i1 and CYCLF-i1 baits. Based on this theory that substitution of the second cysteine in the active site stabilises thioredoxin-target intermediate, if TRX3 or TRX4 could reduce these cysteines, the S42 or S43 mutant is expected to form an intermediate complex with VLCYCLF-i1 or CYCLF-i1, thus prolong or enhance the interaction, which would be observed as an increased level of yeast growth in the RRS. This was tested using the RRS along with the S39, S40, S39S32 and S40S43 mutant preys and the i1, i2, i1-i2, and i1-GGG-i2 baits as controls. Surprisingly, none of the six active site mutants interacted with any of the six baits in the RRS (Figure 6.9). There are three possibilities for this observation. (1)The mutant preys were not expressed in the yeast cells for unknown reasons. Since the prey vector pUra does not contain an epitope tag, we were unable to check prey expression, therefore could not confirm this was the cause of our observation. (2) Both cysteines in the CPPC domain are required for the interaction and substitution of any of the two cysteines would abolish the interaction. As stated earlier, there are cases where substituted active site

mutant abolished the interaction with the target proteins. However, these cases were based on thioredoxin reducing the cysteine residues of the target proteins. Therefore this hypothesis could only explain why the mutants did not interact with VLCYCLF-i1 and CYCLF-i1 which contain cysteines, but could not explain why there was no interaction with the i1, i2 i1-i2 and i1-GGG-i2 baits. (3) It is known that thioredoxin "adopts different conformations in its reduced or oxidized forms and may use protein–protein interactions that depend on a specific conformation as a mechanism for signalling" (Kumar *et al*, 2004). Substitution of the cysteines in the active site may have modified the conformation of TRX3 and TRX4 to an extent that the baits were unable to bind to them anymore.

To investigate whether the absence of interactions in the RRS with the active site mutants was due to a lack of prey expression in yeast, we performed the pull-down assays using i1-GGG-i2-6xHis as bait and Myc-S42 or Myc-S43 as prey. Because the expression of the prey can be monitored by Western blot using the anit-Myc antibody, the pull-down assay would provide a more reliable result than the RRS. The interaction between i1-GGG-i2 and S42 was not detected. This was expected as the interaction between i1-GGG-i2 and TRX3 was not observed in the pull-down. However, as stated in section 6.3.3, i1-GGG-i2 interacts specifically with S43 in the pull-down, which is contrary to the observation that they do not interact in the RRS. The pull-down result supports the view that the absence of interactions in the RRS with the active site mutants was due to a lack of prey expression in yeast. However, we remain cautious in this interpretation, because there are cases that a protein-protein interaction which could not be detected using the yeast-two hybrid can be detected using the pull-down (Kasiviswanathan *et al*., 2005). We can not rule out the possibility that all active site mutants were expressed in yeast and that the absence of interaction between these mutants with the GCR1 derived baits was due to the conformational change of TRX3 and TRX4 caused by the mutation. So far, whether TRX3 and TRX4 reduce the two cysteines in TM1 of GCR1 remains unknown. Performing pull-downs using the S39, S40, S39S40 and S40S43

mutants with the other GCR1 derived baits, especially CYCLF-i1, in the future would provide more information.

In addition to the active site, there is another conserved region in all thioredoxins, which is the flat hydrophobic surface around the active site. The amino acids of this region come from different locations in the primary structure, such as Pro76, Gly92, and Ala93, numbered according to the *E. coli* thioredoxin 1 sequence (Rivera-Madrid *et al*., 1995). This hydrophobic surface is reported to be involved in thioredoxin-target interaction and determining target specificity. For example, a mutant *E.coli* thioredoxin 1 in which Gly92 is substituted by an aspartate residue (hydrophilic, negatively charged) lacks all binding activity to gene 5 protein of phage T7 (Huber *et al*., 1986; Holmgren 1995). Besides, electrostatic bonds have also been reported participating thioredoxin-target interactions. For instance, the L94K/E30K mutation of *E. coli* thioredoxin 1, which has the same redox potential as wild type thioredoxin, was shown to increase binding to chloroplast fructose-1,6-bisphosphatase (FBPase) by 10-fold (Mora-Garcia *et al*. 1998). Likewise, the E63Q and K70Q mutants displayed increased binding ability, whereas the K70E and R74E mutants showed decreased binding ability to FBPase in comparison with that of wild type pea thioredoxin *m* (Wangensteen *et al*., 2001). The NMR structures of human thioredoxin 1 complexed with 13-residue peptide fragments from NF-*κ*B and redox effector factor-1 (Ref-1) provide the first structural evidence of thioredoxin interacting with its targets, which reveals that target proteins interact with thioredoxin at a shallow, crescent-shaped groove in the protein surface, and that the complexes are stabilised by numerous hydrogen bonds, as well as electrostatic and hydrophobic interactions (Yoshioka *et al*., 2006). Further studies of the interactions of GCR1 with site or truncated mutations of TRX3 and TRX4 will facilitate identifying the critical amino acids involved, thus elucidating the nature of the interactions.

In the initially library screening, three thioredoxins – TRX2, TRX3 and TRX4 were identified as potential interactors to GCR1. However, only TRX3 and

TRX4 were confirmed to interact with the i1 and i2 regions in the RRS (section 5.2.5). As all of the three thioredoxins belong to the *Arabidopsis* thioredoxin *h* family, this observation poses an interesting question which is why TRX3 and TRX4, but not TRX2, interact with GCR1. As mentioned in section 6.1.1, TRX2 belongs to subgroup 2, whereas both TRX3 and TRX4 belong to subgroup 1 of the thioredoxin *h* family. A big difference among the three thioredoxins is that both TRX3 and TRX4 contain an atypical active site WCPPC, whereas TRX2 contains the classical active site WCGPC. Such an active site in TRX3 and TRX4 is striking, because proline (P) and glycine (G) have very different structural properties (Brehelin *et al*., 2004). It has been suggested that the presence of the proline in the atypical active site does not change the redox potential of the thioredoxin, but most probably modifies structural details around the active site, and that these structural details may modulate the binding and the reactivity of the thioredoxin with its target (Brehelin *et al*., 2004).

As stated in section 6.1.1, *Brassica* thioredoxins THL1 and THL2, which both contain the CPPC active site, have been shown to interact with the kinase domain of SRK (Bower *et al*., 1996; Mazzurco, *et al* 2001). Of the four *Arabidopsis* thioredoxins *h* proteins tested (TRX1 and TRX2 that contain the GCPC active site, and TRX3 and TRX4 that contain the GPPC active site), and only TRX3 and TRX4 interacted with the kinase domain of SRK (Mazzurco, *et al* 2001). Yeast complementation experiments have demonstrated that only TRX2 can interact with the targets responsible for sulfate assimilation, and only TRX3 and TRX4 would react with the targets implicated in H_2O_2 tolerance (Mouaheb *et al*., 1998; Brehelin *et al*., 2000). Mutation of the TRX3 atypical active site WCPPC to the classical site WCGPC restores a partial sulfate assimilation phenotype (Brehelin *et al*., 2000). Furthermore, mutation of the pea thioredoxin PtTrx*h*3 active site from WCGPC to WCPPC strongly modifies the protein conformation (Gelhaye *et al*., 2003; Gelhaye *et al*., 2005). These data imply that the proline residue of TRX3 and TRX4 might have played a role in the active site conformation, leading to their specific interactions with GCR1. Further studies detailing the interactions of TRX3 and TRX4 WCPPC to

WCGPC mutations with GCR1, as well as TRX2 WCGPC to WCPPC mutation with GCR1 will likely provide valuable information for this hypothesis.

6.3.5 The involvement of TRX3 and TRX4 in GCR1 signalling

ROS, such as H_2O_2 and $O_2^{\cdot-}$, can oxidize cysteine or methionine residues, leading to the formation of disulfide bridges or methionine sulfoxide, and inactivation of cellular proteins (Ritz *et al*., 2000). They are produced in several subcellular compartments in response to environmental stresses and affect many biological processes, including differentiation, transformation, aging, and programmed cell death (Lee *et al*., 2005b). Thioredoxins are known to be involved in the cellular protection against oxidative stress. For instance, thioredoxin is capable of removing H_2O_2, particularly when it is coupled with either thioredoxin peroxidase or methionine sulfoxide reductase (Andoh *et al*., 2002). Mitsui and colleagues (2002) revealed that thioredoxin-overexpressing mice acquired resistance to various oxidative stresses, as well as extended life span. Similarly, thioredoxin has also been reported to play important roles in oxidative stress response and life span regulation in *C. elegans* (Jee *et al*., 2005) and *Drosophila* (Svensson and Larsson, *et al*., 2007). In plants, many enzymes involved in ROS scavenging processes have been isolated as thioredoxin targets, such as ascorbate and secretory peroxidases, dehydroascorbate reductase, and superoxide dismutase (Santos and Rey, 2006). It has also been shown that thioredoxins *h* accumulate in developing wheat seed tissues suffering oxidative stress and are involved in the development of tolerance to oxidative stress during seed desiccation and germination (Serrate and Cejudo, 2003). The induction of thioredoxin is reported to be required for nodule development to reduce ROS levels in soybean roots (Lee *et al*., 2005b). Yeast complementation results that were obtained from various thioredoxins (including AtTRX*h*1-5) expressed in yeast *trx1, trx2* double-mutant cells demonstrated that only TRX3 and TRX4 were able to confer H_2O_2 tolerance (Mouaheb *et al*., 1998; Brehelin *et al*., 2000). In addition, TRX3 has been identified as coupled to 2-Cys peroxiredoxin, a type of

peroxidase involved in H_2O_2 detoxification (Verdoucq *et al*., 1999). Therefore, both TRX3 and TRX4 are likely involved in the oxidative stress response in *Arabidopsis*.

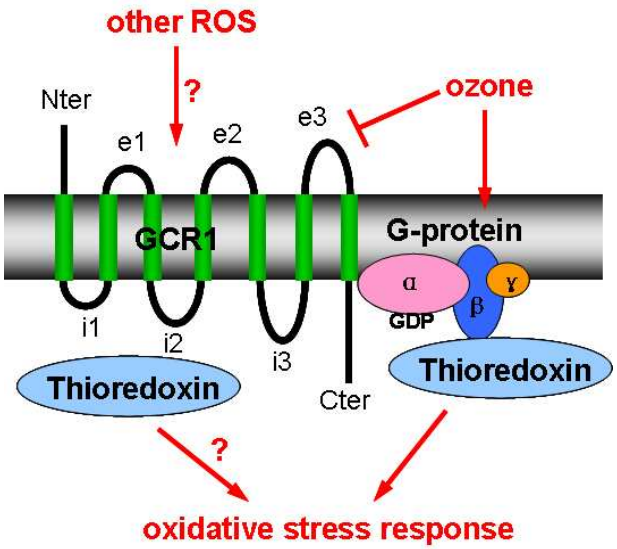


Figure 6.22 A speculative model representing the involvement of thioredoixn, GCR1 and G-protein in oxidative stress response.

As mentioned in section 6.1.1, the G-protein β subunit has been detected as a thioredoxin binding protein (Wong *et al*., 2004, and Alkhalfioui *et al*., 2007). Interestingly, G-protein signalling is required to activate the intracellular source of ROS that contribute to component of the biphasic, stress-elicited oxidative burst in *Arabidopsis* (Booker *et al*., 2004; Joo *et al*., 2005). Joo and colleagues (2005) revealed that the null mutants for Gα and Gβ are less and more sensitive respectively to O_3 damage than wild-type Col-0 plants, and that the early component of the oxidative burst requires both Gα and Gβ whereas the late component requires only Gα. However, the involvement of GCR1 in O_3 elicited stress response was not observed (Booker *et al*., 2004; Joo *et al*., 2005). Nevertheless, the participation of thioredoxins and G-protein in oxidative stress response, together with the GCR1-thioredoxin, GCR1-G-protein and thioredoxin-G-Protein interactions, lead us to postulate that thioredoxin may function as a key regulator that modifies GCR1 and G-protein activities in

response to various oxidative stress conditions, and that the GCR1-thioredoxin interaction might be part of the signalling pathway in oxidative stress response elicited by other ROS, such as H_2O_2 (Figure 6.22). Comparing the response of the *gcr1 trx3*, *gcr1 trx4*, *gcr1 trx3 trx4* double or triple mutants to H_2O_2 and other ROS elicited oxidative stresses with those of the *gcr1*, *trx3* and *trx4* single mutants and wild type *Arabidopsis* plants, would be of great interest.

Accumulating evidence implicates that in addition to being toxicants, ROS also act as essential signalling molecules to regulate biological processes, such as growth, cell cycle, hormone signalling, stress response and development (Mittler *et al*, 2004; Misra *et al*., 2007). It has been suggested that they trigger signalling pathways in particularly via thioredoxins (Felberbaum-Cort *et al*., 2007). For examples, the thioredoxin-like protein TXNL1 can interact with the Guanine nucleotide Dissociation Inhibitor (GDI), an interactor of the small GTPase Rab5 (Felberbaum-Cort *et al*., 2007). TXNL1 functions as "an effector of ROS or a redox sensor by converting redox changes into changes of GDI capacity to capture Rab5, which in turn modulates fluid phase endocytosis" (Felberbaum-Cort *et al*., 2007). Interestingly, a few GPCR agonists, e.g. angiotensin II, serotonin [5-hydroxytryptamine (5-HT)], thrombin and endothelin-1 have been shown to generate ROS in different cell systems in mammals (Thannickal and Fanburg, 2000). Furthermore, GPCRs can activate JAK/STAT (Janus tyrosine kinases/signal transducers and activators of transcription) signalling through a Rac (small G-proteins)–dependent generation of ROS (Pelletier *et al*., 2003), suggesting a role for GPCR in mediating ROS-related signal transduction. Therefore, it is also possible that thioredoxin acts as a redox sensor to regulate GCR1-mediated ROS-related signalling pathways.

Molecular chaperones form a class of polypeptide-binding proteins that are implicated in protein folding, protein targeting to membranes, protein renaturation or degradation after stress and the control of protein-protein interactions (Kern *et al*., 2003). Some GPCR-interacting proteins have been

shown to behave as chaperones and are involved in surface expression of the newly synthesized or recycled GPCR (Brady and Limbird, 2002; Ulloa-Aguirre *et al*., 2004). For instance, in *Drosophila melanogaster*, the absence of the chaperone Nina A (*n*either *i*nactivation *n*or *a*fterpotential A) leads to rhodopsin endoplasmic reticulum (ER) accumulation and ultimately to its degradation (Colley *et al*., 1991). RAMPs is a family of chaperone proteins, which has been found to not only assist in the transport of calcitonin receptor-like receptor to the cell surface, but also define the glycosylation state and recognition properties of the receptor (McLatchie *et al*., 1998). Another GPCR-interacting protein, Calnexin, is a molecular chaperone that recognizes immature glycosylation states of membrane proteins and assists in the folding of newly synthesized proteins or, alternatively, targets improperly folded molecules to the degradation pathway (Brady and Limbird, 2002). Recently, Calnexin has also been found to facilitate the formation of GPCR dimers prior to their trafficking to the cell surface (Free *et al*., 2007).

An increasing body of evidence indicates that thioredoxin also functions as chaperone and its chaperone activity is independent of its role in the catalysis of disulfide bond reactions. For instance, an active site mutant C35A of *E.coli* thioredoxin 1 did not interfere with its chaperone activity (Kern *et al*., 2003). Similarly, *E.coli* thioredoxin 1 CGPC to AGPA mutant promotes the correct folding of single chain Fv antibodies (scFvs), indicating that thioredoxin 1 acts largely as an intramolecular protein chaperone in the correct folding of scFvs, but not as a disulfide bond catalyst (Jurado *et al*., 2006). Our observation that TRX3 and TRX4 interact with all parts of GCR1 in the rRRS (6.3.2) indicates that they could possibly function as chaperones to regulate GCR1 activities, such as assisting the correct folding of newly synthesized protein, targeting it to the cell membrane, or targeting misfolded GCR1 to its degradation pathway.

Chapter 7 Characterisation of the interaction between GCR1 and a DHHC zinc finger protein *zf*-DHHC1

7.1 Introduction

In addition to TRX3 and TRX4, the other potential interactor that had been verified to interact with the i1 and i2 regions of GCR1 (Chapter 5) was an Asp-His-His-Cys (DHHC) type zinc finger family protein (referred to henceforth as *zf*-DHHC1). It contains the pfam01529 DHHC zinc finger domain, which is a cysteine rich domain that is highly conserved in organisms from yeast to human (Li *et al*, 2002). Genes encoding DHHC proteins are found in all eukaryotes examined to date, with the number of examples ranging from 1-8 in unicellular fungi, to more than 20 in metazoans (Mitchell *et al*., 2006).

Newly synthesized proteins can be subject to post-translational lipid modifications, such as myristoylation, prenylation and palmitoylation, which serve to tether them to the cytoplasmic surface of cellular membrane. A growing body of data suggests that proteins with the DHHC domain are palmitoyl acyl transferases (PATs), which is also known as palmitoyltransferases, that catalyze the attachment of palmitate, a 16-carbon saturated fatty acid, to cysteine residues on substrate proteins (Gleason *et al*., 2006; Resh 2006). For instance, the Golgi-specific DHHC protein GODZ palmitoylates the γ2 subunit of the $GABA_A$ receptor in mouse (Keller *et al*., 2004). The yeast DHHC protein Swf1 is required for the palmitoylation of the soluble N-ethylmaleimide-sensitive fusion protein attachment protein receptor (SNARE) Snc1, Syn8 and Tlg1 (Valdez-Taubas and Pelham, 2005). The Erf2p-Erf4p complex is required for the catalytic transfer of palmitate to yeast Ras2p, and mutations within the conserved residues of the Erf2p DHHC domain abolish Ras2p palmitoylation (Lobo *et al*., 2002). A recent palmitoylproteomics analysis of yeast strains, in which the genes encoding DHHC PATs were

deleted individually or in combination, has revealed that these enzymes account for most of the cellular palmitoylation events in yeast (Roth *et al*., 2006).

Palmitoylation differs from myristoylation and prenylation, in that it is often found to modify integral membrane proteins, thus the list of palmitoylated proteins also includes many GPCRs (Roth *et al*, 2006), such as bovine rhodopsin (Ovchinnikov *et al*., 1988), human β_2-adrenergic receptor (O'Dowd *et al*., 1989), canine H2 histamine receptor (Fukushima *et al*., 2001) and human CCR5 chemokine receptor (Percherancier *et al*., 2001). These GPCRs have been shown to be palmitoylated at cysteines within the C-terminal tail. It is believed that the covalent bound palmitate become intercalated in the membrane bilayer, thereby creating a fourth cytoplasmic loop (Watson and Arkinstall, 1994). The crystal structure of rhodopsin provided structural evidence that palmitoylation indeed results in the formation of a fourth cytoplasmic loop (Palczewski *et al*., 2000). Cysteine residues at similar locations are found in about 80% of all GPCRs, indicating that palmitoylation is a general characteristic of this type of receptor (Escriba *et al*., 2007).

7.1.1 Research objectives and experimental approach

Most DHHC PATs are integral membrane proteins (Mitchell *et al*., 2006) with four TM domains (in some cases two, three, five or six) (Valdez-Taubas and Pelham, 2005), and the DHHC domain is generally located between TM2 and TM3, extending into TM3 (Smotrys and Linder 2004; Keller *et al*., 2004; figure 7.1).

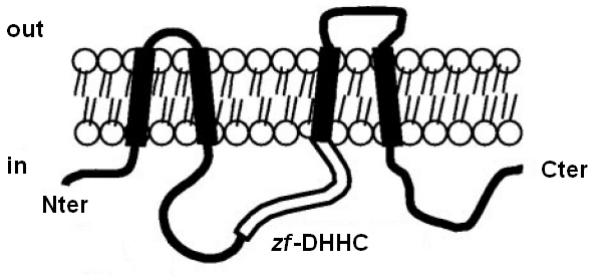


Figure 7.1 Proposed transmembrane topography of GODZ (taken from Keller *et al*., 2004)

It was our great interest to find out whether *zf*-DHHC1 possesses these features shared by the other DHHC PATs, because to some extent, functional information can be derived from structural and sequential similarities. Therefore the first objective was to predict the membrane topography of *zf*-DHHC1 using three online programs TMpred, TMHMM and TopPred (section 2.2.1). The amino acid sequence of *zf*-DHHC1 would also be aligned with two well studied yeast DHHC PATs, Swf1 (Valdez-Taubas and Pelham, 2005) and Erf2p (Lobo *et al*., 2002), using an online multiple sequence alignment programme ClustalW (2.2.18).

Many GPCRs are palmitoylated at the cysteine residues within their C-terminal tails. However, there is an increasing body of evidence that points to the presence of palmitoylated cysteines outside this receptor domain (Qanbar and Bouvier, 2003). In fact, the position of palmitoylated cysteines varies considerably in different palmitoylated proteins (Ponimaskin and Schmidt, 1998). Some integral membrane proteins are palmitoylated within their intracellular regions while many others are palmitoylated near the membrane/cytoplasm border in the TM domain. It has been suggested that any cysteine that is a few residues from the membrane/cytoplasm border in the TM region can be palmitoylated (Ochsenbauer-Jambor, *et al*., 2001). Although there is no cysteine present in its C-terminal tail and the three intracellular loops of GCR1, there are two cysteines in the VLCYCLF motif in the first TM domain near i1, which could be a potential palmitoylation site for *zf*-DHHC1. Therefore, the second objective was to investigate whether *zf*-DHHC1 interacts with the motif containing the two cysteines by testing whether its interaction with the VLCYCLF-i1 and CYCLF-i1 baits would be greater than with the i1 bait in the RRS. This is based on the assumption that if *zf*-DHHC1 interacts with these cysteines, the addition of these cysteines to the i1 bait might enhance the interaction between *zf*-DHHC1 and i1.

The third objective was to verify the interaction between GCR1 and *zf*-DHHC1 using the rRRS. The *zf*-DHHC1 membrane topography prediction result would

be useful for this part of experiment as well, because if *zf*-DHHC1 was predicted to be an integral membrane protein, it would have to be used as the bait in the rRRS and would be expected to be localised to the cell membrane. Otherwise, when used as a prey and fused to the mRas, it would bring the mRas to the cell membrane and autoactivate the system.

The finally objective was to verify the GCR1-*zf*-DHHC1 interaction using the pull-down method, in the same way as we performed for the GCR1-TRX3 and GCR1-TRX4 interactions (section 6.3.4).

7.2 Results

7.2.1 *zf*-DHHC1 transmembrane topography prediction

The transmembrane topography of *zf*-DHHC1 was predicted using three protein transmembrane topography prediction programs: TMpred, TMHMM and TopPred (Table 7.1).

Table 7.1 Transmembrane topography prediction for *zf*-DHHC1

TM	TM regions predicted by different programs					Most likely TM regions
	TMpred			TMHMM	TopPred	
	17aa*	19aa*	21aa*			
TM1	34-53	34-53	34-54	34-56	36-56	34-54
TM2	66-85	66-85	66-87	66-88	66-86	66-86
TM3	207-232	207-232	207-232	207-226	214-234	207-232
TM4	240-262	240-262	240-262	241-263	242-262	241-262

*Three TMpred predictions were made using 17aa, 19aa and 21aa as the minimum length of hydrophobic helix respectively. Default settings were used for TMHMM and TopPred predictions. Numbers indicate the amino acid residues at the start and end of each TM region. The final column (shaded) shows the most likely positions of TM regions based on this analysis.

All three programs predict *zf*-DHHC1 to be an integral membrane protein, with a cytoplasmic Nter, four TM domains linked by one intracellular loop and two

extracellular loops, and a cytoplasmic Cter. The DHHC domain (64 aa long, from aa153 to aa217) is located between the TM2 and TM3 domains, extending into TM3 (Figure 7.2), which fits very well in the structure model of a typical DHHC PAT. The multiple amino acid sequence alignment for *zf*-DHHC1 with two well studied yeast DHHC PATs, Swf1 and Erf2p, was performed using ClustalW (section 2.2.18). As shown in figure 7.2, the greatest homology is around the regions of the DHHC domain, and conserved amino acid residues within the DHHC domains of Swf1 and Erf2p are also present in *zf*-DHHC1.

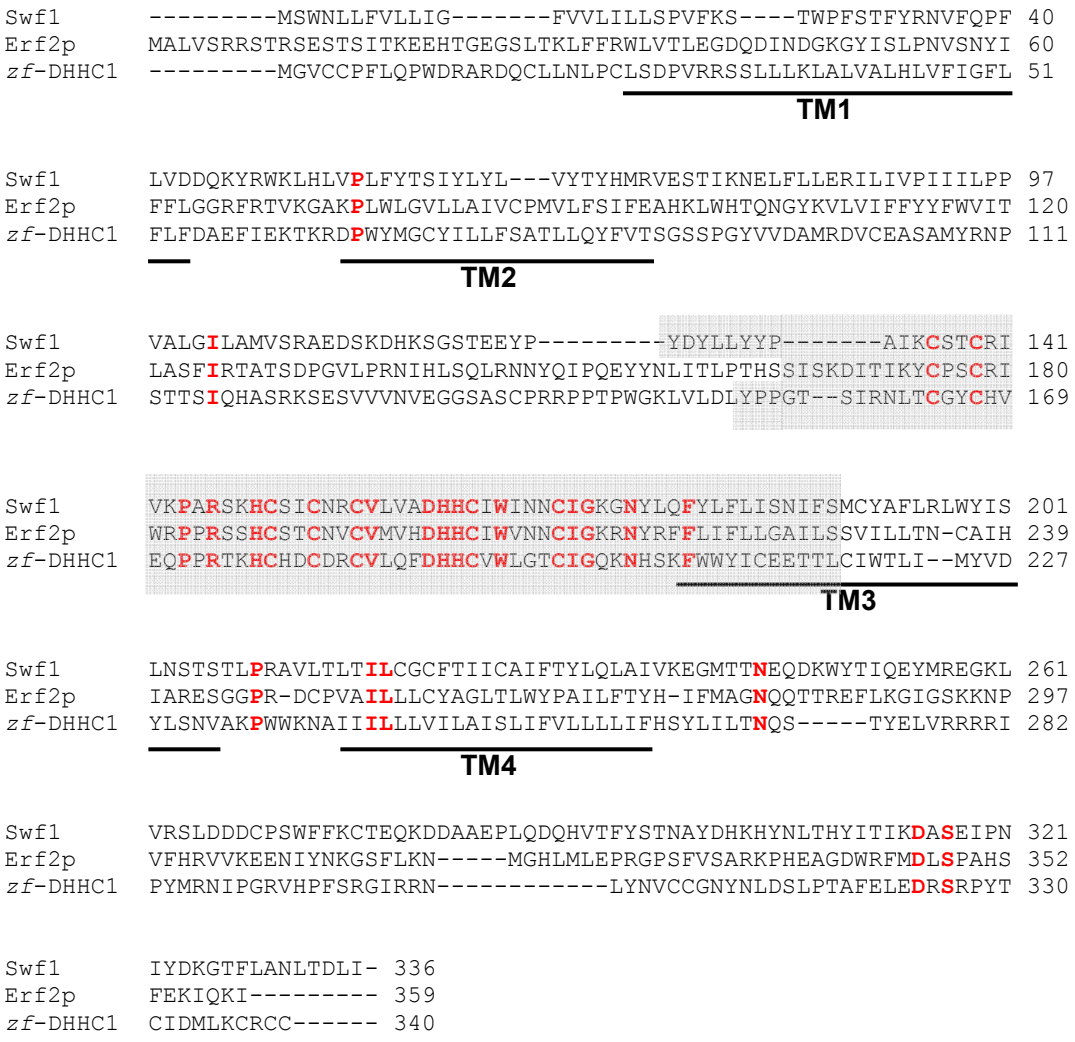
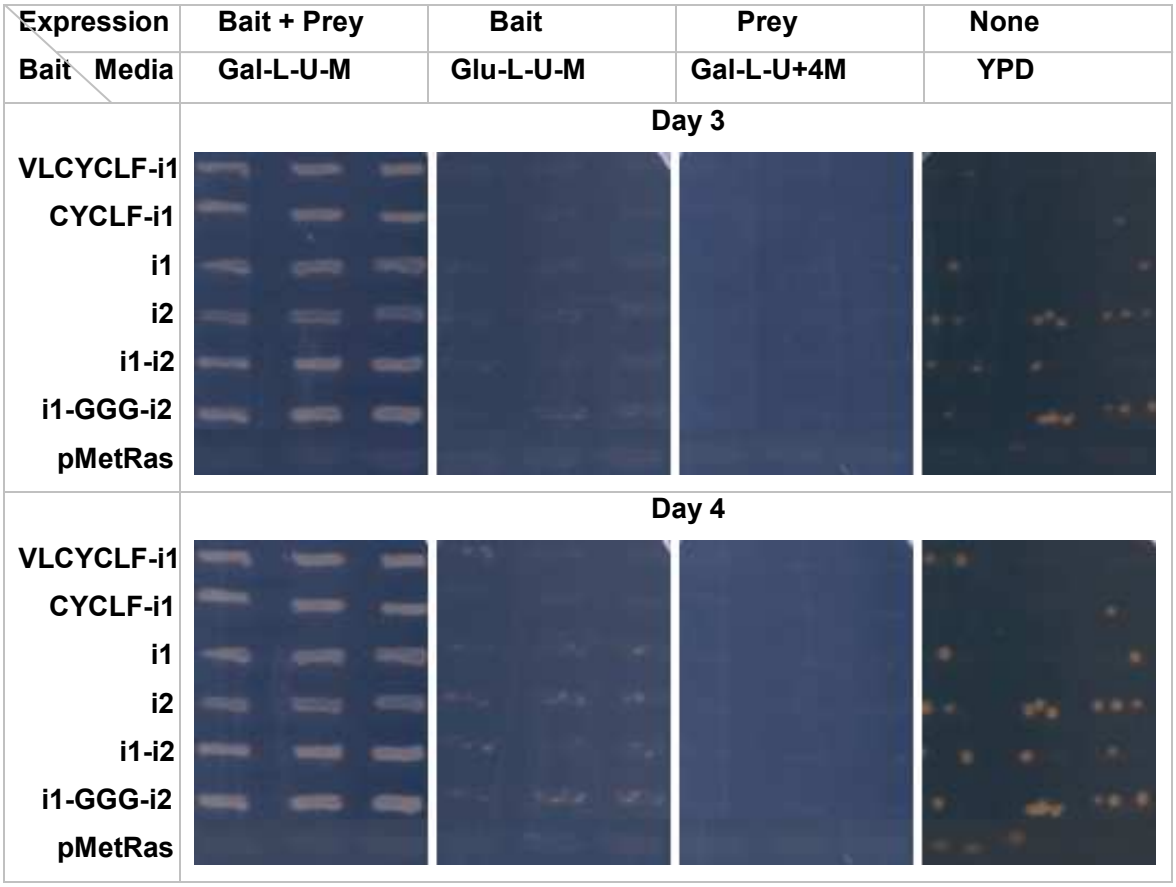


Figure 7.2 Multiple sequence alignment for *zf*-DHHC1, Swf1 and Erf2p, and proposed transmembrane topography of *zf*-DHHC1. Putative TM domains of *zf*-DHHC1 are underlined. The DHHC domain region of each sequence is shaded in gray. Residues that are identical in the three sequences are indicated by red and bold letters.

7.2.2 The addition of cysteines to the i1 bait enhances its interaction with *zf*-DHHC1 in the RRS

To examine whether *zf*-DHHC1 interacts with VLCYCLF-i1 and CYCLF-i1, the pUra-*zf*-DHHC1 construct was transformed into the bait-containing (VLCYCLF-i1 and CYCLF-i1) yeast cells and were plated onto Glu-L-U+M medium and incubated at 24°C until colonies appeared. The colonies were streaked onto the Glu-L-U+M plates and incubated at 24°C until they had grown sufficiently to be replica plated. They were then replica plated onto the selective medium Gal-L-U-M and the three control media Glu-L-U-M, Gal-L-U+4M and YPD, and incubated at 36°C for 6 days. The interactions for *zf*-DHHC1 with the i1, i2, i1-i2 and i1-GGG-i2 baits and the polypeptide encoded by the empty bait vector pMetRas were included in the experiment as controls. The plates were scanned every 24 hours from day 3 until day 6 (Figure 7.3).



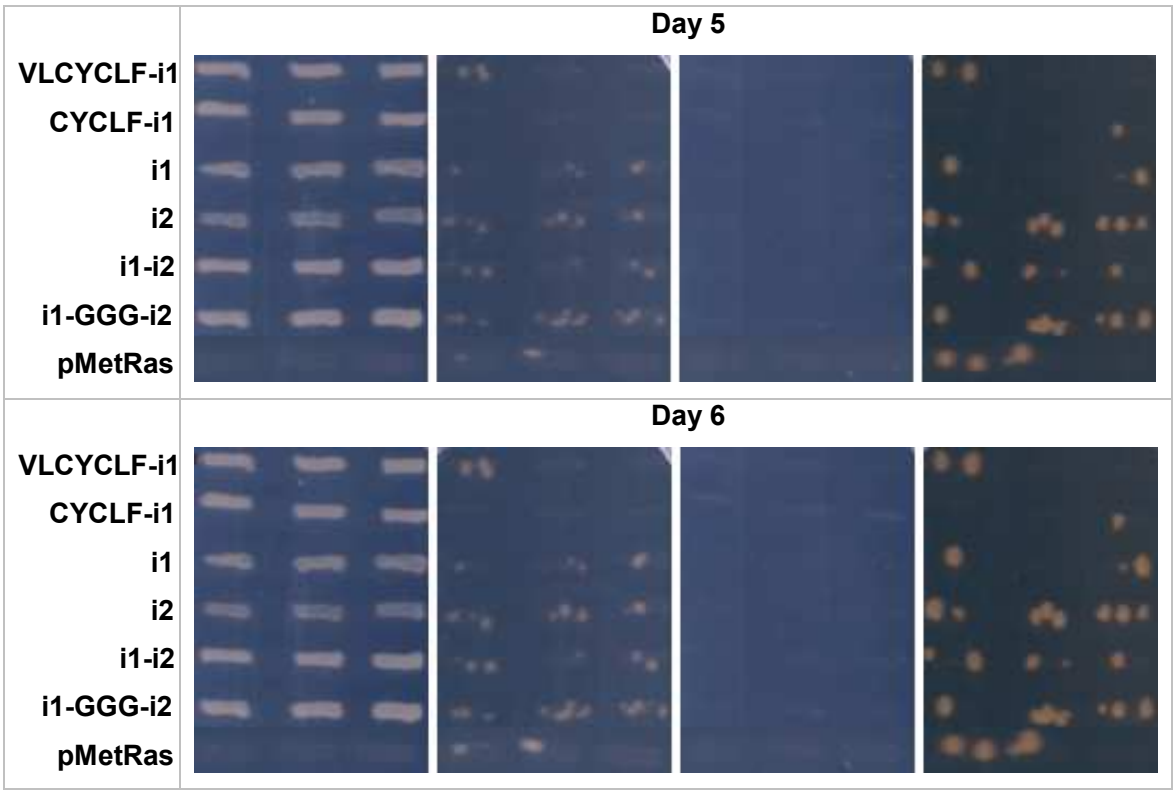


Figure 7.3 Interactions between the GCR1 derived baits and the *zf*-DHHC1 prey. The yeast streaks were replica plated onto the four media/plates and incubated at 36°C for 6 days. The plates were scanned every 24 hours from day 3 to day 6.

As can be seen on the Gal-L-U-M medium (Figure 7.3), on which both bait and prey were expressed, there was definite growth for all interactions between *zf*-DHHC1 and the GCR1 derived baits, but there was no growth for *zf*-DHHC1 with the polypeptide encoded by empty bait vector pMetRas. This result indicates that *zf*-DHHC1 interact with i1, i2, and the two extended i1 baits VLCYCLF-i1 and CYCLF-i1, and that the interactions seen were not caused by prey interacting with the polypeptide encoded by empty bait vector. Furthermore, yeast containing i1 and i2, especially i2, had lower level of growth, and all the others (VLCYCLF-i1, CYCLF-i1, i1-i2, and i1-GGG-i2) had the similar higher level of growth. Because both VLCYCLF-i1 and CYCLF-i1 are expressed at a much lower level than the other baits (Figure 6.4), interaction for *zf*-DHHC1 with VLCYCLF-i1 and CYCLF-i1 could well much greater than with i1, i2, i1-i2, and i1-GGG-i2. There were three transformants for each bait-prey

interaction and the interactions seen for the three transformants were identical, thus the result was reproducible.

7.2.3 GCR1 interacts with *zf*-DHHC1 in the rRRS

In order to reconfirm the GCR1-*zf*-DHHC1 interaction seen in the RRS, the interactions between *zf*-DHHC1 and the GCR1 derived baits were also examined in the rRRS. Because *zf*-DHHC1 was predicted to be an integral membrane protein (section 7.2.1), it was used as the bait, whereas regions of GCR1 (VLCYCLF-i1, CYCLF-i1, i1, i2, i1-i2 and i1-GGG-i2) were used as the preys in the rRRS.

7.2.3.1 Making the pMet-*zf*-DHHC1 rRRS bait construct

The *zf*-DHHC1 insert was PCR amplified using the Expand High Fidelity^{PLUS} PCR System (section 2.2.2.1), with the *zf*-DHHC1 cDNA as template and the ZFBegHindIII + ZFEnd primer pair (Table 2.6). The predicted size of the insert is 1032 bp, and the PCR product obtained was consistent with this (Figure 7.4). The PCR product was initially cloned into pCR2.1 using the TOPO cloning method as described in section 2.2.1.6. Recombinants were checked for insert size and orientation by colony PCR with a vector-specific M13F as forward primer and an insert-specific ZFEnd as reverse primer (data not shown). Positive clone was identified and the plasmid pCR2.1-*zf*-DHHC1 was extracted.

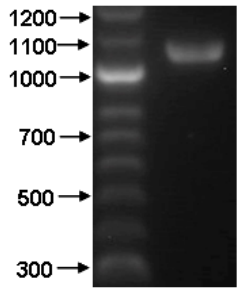


Figure 7.4 PCR amplification of the insert. The *zf*-DHHC1 insert was PCR amplified. The predicted size of the insert is 1032 bp.

The pCR2.1-*zf*-DHHC1 plasmid was then digested with *HindIII*, and the fragment of expected size for *zf*-DHHC1 was gel-purified and ligated into the *HindIII* digested rRRS bait expression vector pMet. Recombinants were checked for insert size and orientation by colony PCR with a vector-specific pMetF as forward primer and an insert-specific ZFEnd as reverse primer (data not shown). Positive clone was identified and the plasmid was extracted and sequenced (data not shown). Sequencing result confirmed that *zf*-DHHC1 was successfully cloned in frame into the *HindIII* site in pMet, which generated the pMet-*zf*-DHHC1 construct.

7.2.3.2 Making the pUraRas-VLCYCLF-i1, pUraRas-CYCLF-i1, pUraRas-i1, pUraRas-i2, pUraRas-i1-i2 and pUraRas-i1-GGG-i2 rRRS prey constructs

The VLCYCLF-i1, CYCLF-i1, i1, i2, i1-i2 and i1-GGG-i2 inserts were PCR amplified using the Expand High Fidelity^{PLUS} PCR System (section 2.2.2.1). They were amplified using the corresponding RRS constructs as templates, with the VLP1FEcoRI + LP1EndEcoRI, CLP1FEcoRI + LP1EndEcoRI, LP1FEcoRI + LP1EndEcoRI, LP2FEcoRI + LP2EndEcoRI, LP1FEcoRI + LP2EndEcoRI and LP1FEcoRI + LP2EndEcoRI primer pairs respectively (Table 2.6). The predicted size of each insert is VLCYCLF-i1: 80 bp, CYCLF-i1: 74 bp, i1: 59 bp, i2: 77 bp, i1-i2: 104 bp and i1-GGG-i2: 113 bp, and the PCR products obtained were consistent with this (Figure 7.5).

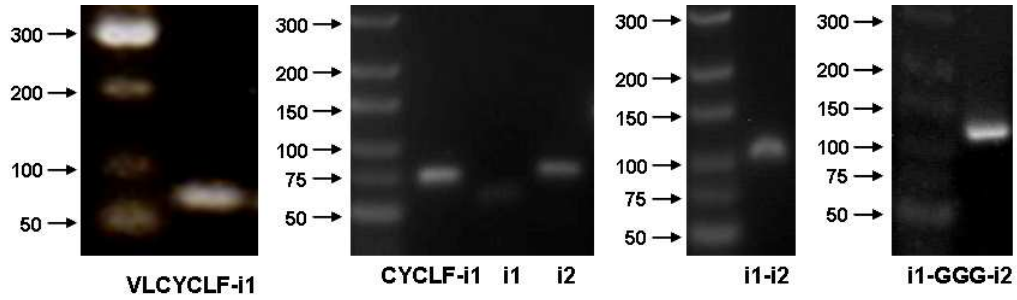


Figure 7.5 PCR amplification of the inserts. The six inserts VLCYCLF-i1, CYCLF-i1, i1, i2, i1-i2 and i1-GGG-i2 were PCR amplified. The predicted size of each insert is: 80 bp, 74 bp, 59 bp, 77 bp, 104 bp and 113 bp respectively.

The PCR products were digested with *EcoRI* and ligated into the *EcoRI* digested rRRS prey expression vector pUraRas. Recombinants were checked for insert size and orientation by colony PCR with a vector-specific T7 as forward primer, and an insert-specific primer e.g. LP1EndEcoRI as reverse primer (data not shown). Positive clones were identified and the plasmids were extracted and sequenced (data not shown). Sequencing results confirmed that all inserts were successfully cloned in frame to the C-terminus of Ras in pUraRas, which generated the pUraRas-VLCYCLF-i1, pUraRas-CYCLF-i1, pUraRas-i1, pUraRas-i2, pUraRas-i1-i2 and pUraRas-i1-GGG-i2 constructs.

7.2.3.3 Autoactivation test of the VLCYCLF-i1, CYCLF-i1, i1, i2, i1-i2, i1-GGG-i2 rRRS preys

The pUraRas-VLCYCLF-i1, pUraRas-CYCLF-i1, pUraRas-i1, pUraRas-i2, pUraRas-i1-i2 and pUraRas-i1-GGG-i2 constructs were transformed into the temperature sensitive yeast strain *cdc25-2*. Four yeast colonies for each prey were randomly selected for the temperature sensitivity test. They were streaked onto a Glu-U+M plate and incubated at 24°C for 3 days. They were then replica plated onto two Glu-U+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as a control. All yeast streaks grew at 24°C, but there was virtually no yeast growth at 36°C (Figure 7.6). Therefore these prey-containing yeast cells were temperature sensitive and were used in the following experiments.

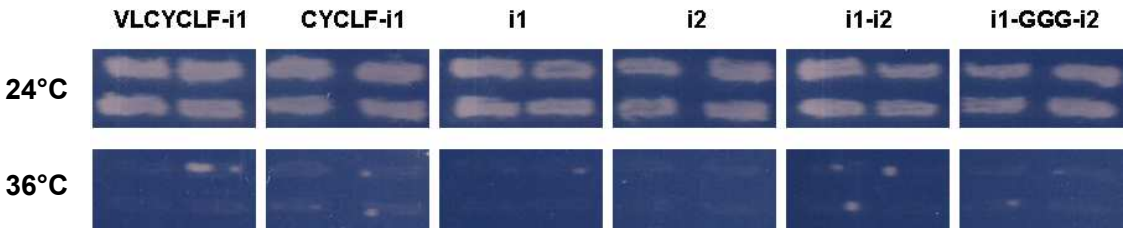


Figure 7.6 Prey temperature sensitivity test. The prey-containing (VLCYCLF-i1, CYCLF-i1, i1, i2, i1-i2 and i1-GGG-i2) yeast cells were streaked onto Glu-U+M plate and incubated at 24°C for 3 days. They were then replica plated onto two Glu-U+M plates, one was grown at 36°C for 3 days, and the other at 24°C for 3 days as a control. Photos show yeast growth after this 3 day period.

The rRRS prey vector pUraRas contains mRas, which might be brought to the cell membrane when it is fused to the prey, resulting in autoactivation of the system. It is vital to work only with preys that do not autoactivate the rRRS. To check for autoactivation, yeast cells expressing prey with no bait were replica plated onto two Gal-U+M plates, one was grown at 36°C for 5 days, and the other at 24°C for 5 days as a control. Yeast cells containing the VLCYCLF-i1, CYCLF-i1 and i2 preys grew at 24°C, but there was virtually no yeast growth at 36°C (Figure 7.7), indicating that these preys did not autoactivate the rRRS, hence can be used in the following rRRS screens. However, there was yeast growth at both 24°C and 36°C for the i1, i1-i2 and i1-GGG-i2 preys (Figure 7.7), indicating that they did autoactivate the rRRS, and were therefore not suitable for use in the following rRRS screens.

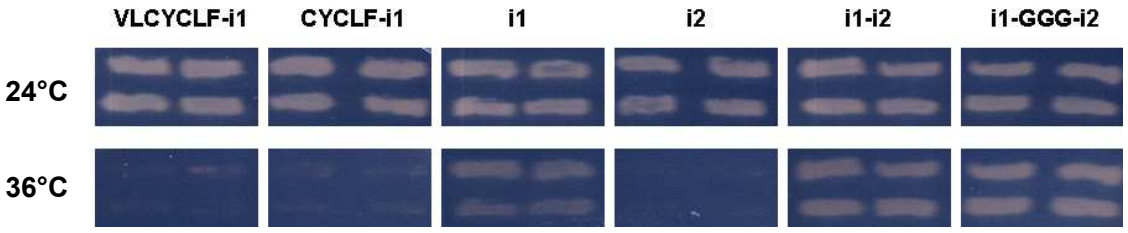
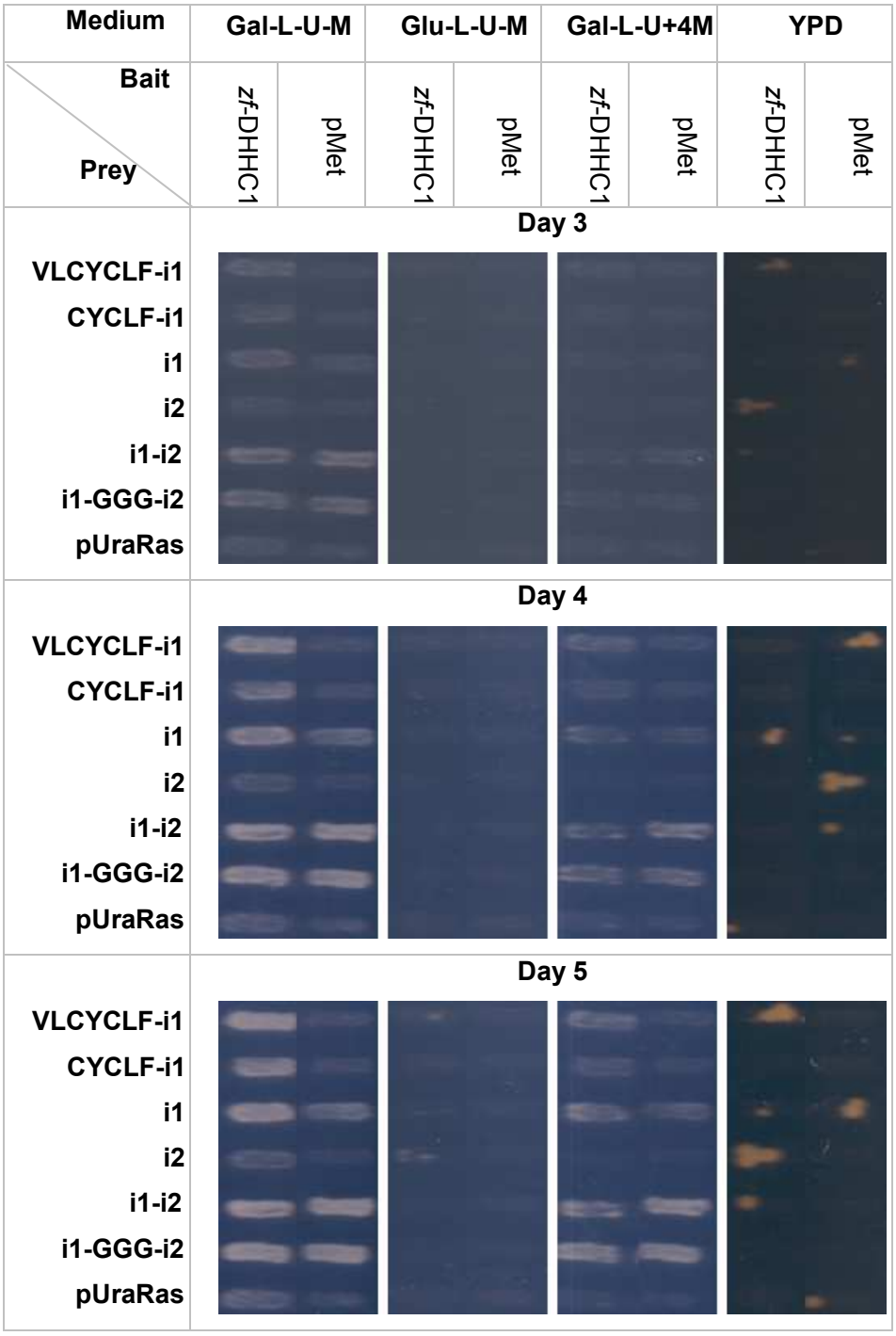


Figure 7.7 Prey autoactivation test. To check whether the preys autoactivate the rRRS or not, the yeast cells were replica plated onto two Gal-U+M plates, one was grown at 36°C for 5 days, and the other at 24 °C for 5 days as a control. Photos show yeast growth after this 5 day period.

7.2.3.4 Interactions between the *zf*-DHHC1 bait and the GCR1 derived preys in the rRRS

The pUraRas-VLCYCLF-i1, pUraRas-CYCLF-i1, pUraRas-i1, pUraRas-i2, pUraRas-i1-i2, pUraRas-i1-GGG-i2 prey constructs and the empty prey vector pUraRas were co-transformed separately with the pMet-*zf*-DHHC1 bait construct into the temperature sensitive yeast strain *cdc25-2*. In parallel, the empty bait vector pMet was co-transformed with the prey constructs, and were used as controls. The yeast cells were plated onto Glu-L-U+M plates and incubated at 24°C until colonies appeared. Randomly selected colonies for each of the bait-prey combination were streaked onto the Glu-L-U+M plates

and incubated at 24°C until they had grown sufficiently to be replica plated. They were then replica plated onto the selective medium Gal-L-U-M and the three control media Glu-L-U-M, Gal-L-U+4M and YPD and incubated at 36°C for 6 days. The plates were scanned every 24 hours from day 3 until day 6 (Figure 7.8).



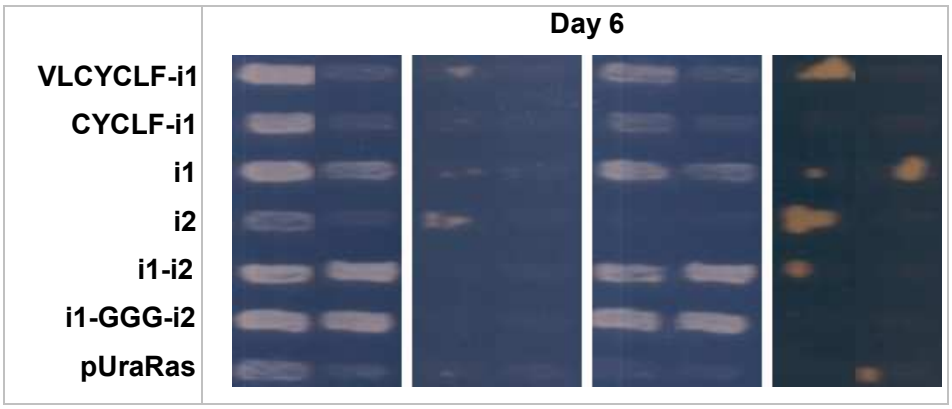


Figure 7.8 Interactions between the *zf*-DHHC1 bait and the GCR1 derived preys in the rRRS. The yeast streaks were replica plated onto the four media/plates and incubated at 36°C for 6 days. The plates were scanned every 24 hours from day 3 to day 6.

As shown in Figure 7.8, over the course of the experiment, there was a much higher level of yeast growth for *zf*-DHHC1 with VLCYCLF-i1 and CYCLF-i1 than with i2 and pUraRas on the Gal-L-U-M plate. In addition, there was no yeast growth for the control group that composed of pMet with VLCYCLF-i1, CYCLF-i1, i2 and pUraRas. These results indicate that there was interaction between the VLCYCLF-i1, CYCLF-i1 preys and the *zf*-DHHC1 bait, and the interaction was not caused by prey interacting with the polypeptide encoded by the empty bait vector. There was a very low level of growth for both i2 and pUraRas, and difference in their growth level was not particularly discernable, which suggests that there was either no interaction between *zf*-DHHC1 and i2, or there was only very weak interaction. In addition, there was yeast growth on the Gal-L-U-M media for the i1, i1-i2 and i1-GGG-i2 preys with not only the *zf*-DHHC1 bait but also the polypeptide encoded by the empty bait vector pMet. Since these preys show autoactivation, this growth pattern was expected.

7.2.4 Cloning and expression of Myc-*zf*-DHHC1 in *E.coli*

The pCR2.1-*zf*-DHHC1 plasmid was digested with *HindIII*, and the fragment of the expected size for *zf*-DHHC1 (1032 bp) was gel-purified and ligated into the *HindIII* digested *E.coli* expression vector pET22b(+)myc (Table 2.2, figure 7.9). Recombinants were checked for insert size and orientation by colony PCR with

a vector-specific T7 as forward primer and an insert-specific ZFEnd as reverse primer (data not shown). Positive clone was identified and the plasmid was extracted and sequenced (data not shown). Sequencing result confirmed that the insert was successfully cloned in frame to the C-terminus of Myc in pET22b(+)myc (*HindIII* site), which generated the pET22b(+)myc-*zf*-DHHC1 construct.

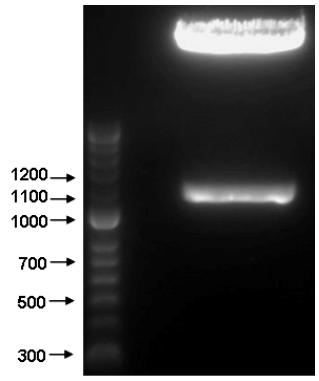


Figure 7.9 Preparation of the insert. The pCR2.1-*zf*-DHHC1 plasmid was digested with *HindIII*, and the fragment of the expected size for *zf*-DHHC1 (1032 bp) was gel-purified.

The pET22b(+)myc-*zf*-DHHC1 construct was transformed into the *E.coli* expression strain BL21 (DE3), and the expression of the recombinant protein was checked as described in section 2.2.13. Although a few attempts have been made, including using different protein expression induction conditions and various *E.coli* strains, the expression of *zf*-DHHC1 was not detected (data not shown).

7.3 Discussion

7.3.1 *zf*-DHHC1 interacts with the extended intracellular loop 1 baits containing two cysteines from TM1 of GCR1

The *zf*-DHHC1 was identified in the initial library screening as a GCR1 interactor (chapter 4), and was subsequently confirmed to interact with the i1 and i2 regions of GCR1 in the RRS (chapter 5). DHHC proteins are known to

palmitoylate target protein on the cysteine residues, but neither of the two loops contains cysteines. However, there are two cysteines in the motif VLCYCLF in the TM1 region near i1 (Figure 3.1). It has been suggested that any cysteine that is a few residues from the membrane/cytoplasm border in the TM region can be palmitoylated (Ochsenbauer-Jambor, *et al*., 2001), hence the two cysteines in TM1 of GCR1 could be potential palmitoylaiton sites for *zf*-DHHC1. If *zf*-DHHC1 interacts with these two cysteines in GCR1, the addition of these cysteines to the i1 bait may enhance the interaction between i1 and *zf*-DHHC1. Therefore, the interaction between *zf*-DHHC1 and two extended i1 baits (VLCYCLF-i1 and CYCLF-i1) containing the two cysteines were examined using the RRS, and the strength of their interaction was compared with the interaction between *zf*-DHHC1 and i1.

As stated in section 6.3.1, the stronger the interaction between bait and prey, the more likely it is that the mRas will be membrane localised quickly and retained long enough to induce cell proliferation. It is therefore assumed that a high level of yeast growth implicates a strong degree of interaction between the bait and prey. If the cysteine residues are important for the interaction then the strength of the interactions for *zf*-DHHC1 with the VLCYCLF-i1 and CYCLF-i1 baits is expected to be greater than with the i1 bait. As shown in figure 7.3, yeast growth was higher for VLCYCLF-i1 and CYCLF-i1 compared with i1, suggesting a greater interaction. However, the Western blot analysis shows that the baits are not expressed at the same level (Figure 6.4) so the differences in yeast growth seen may not be due entirely to the strength of the protein-protein interaction. In fact, both VLCYCLF-i1 and CYCLF-i1 are expressed at a much lower level than i1 (Figure 6.4), thus the interaction for *zf*-DHHC1 with VLCYCLF-i1 and CYCLF-i1 could well much greater than with i1. Another possibility should be considered is that the level of interaction may be limited by the amount of prey present in the yeast cells, which could lead to saturation of the growth response. Given that the level of yeast growth varied for different baits, it is more likely that there was sufficient prey, and that the level of yeast growth was a reflection of interaction strength for individual bait.

Altogether, *zf*-DHHC1 did have a stronger interaction with VLCYCLF-i1 and CYCLF-i1 compared with i1, indicating that it interacts strongly with the region containing the two cysteines in TM1 of GCR1.

In addition, as shown in figure 7.3, i2 has the weakest interaction with *zf*-DHHC1 compared with other baits. The observation that i2 was expressed at a much higher level than i1 (Figure 6.4), and that i2 had a lower level of yeast growth than i1 (Figure 7.3) indicates that *zf*-DHHC1 binds preferentially to i1 than i2 during a GCR1 interaction. Furthermore, the interaction may have been enhanced when both i1 and i2 were present, since the level of yeast growth for i1-i2 and i1-GGG-i2 were higher than for i1 and i2 (Figure 7.3). However, it is also possible that the lower level of yeast growth for i1 compared with i1-i2 and i1-GGG-i2 was due to its lower expression level than i1-i2 and i1-GGG-i2 (Figure 6.4).

7.3.2 GCR1 interacts with *zf*-DHHC1 in the rRRS

The rRRS screens were performed in order to reconfirm the GCR1-*zf*-DHHC1 interaction seen in the RRS. Originally, it was intended to use *zf*-DHHC1 as the prey and different regions of GCR1 (Nter-i2, i3-Cter and Nter-Cter) as the baits to examine whether *zf*-DHHC1 interacts with the other regions of GCR1. However, *zf*-DHHC1 was predicted to be an integral membrane protein (section 7.2.1), so it is not suitable for use as the prey, or else it would bring the fused mRas to the cell membrane and autoactivate the rRRS. Therefore *zf*-DHHC1 was used as the bait, whereas VLCYCLF-i1, CYCLF-i1, i1, i2, i1-i2 and i1-GGG-i2 were used as the preys (section 7.2.3).

As shown in figure 7.7, the i1, i1-i2 and i1-GGG-i2 preys showed autoactivatinon in the rRRS. The fact that the autoactivation occurred is unexpected as none showed autoactivation in the RRS. The possible reason could be that instead of the bait-Myc-mRas order for the RRS vector pMetRas, the rRRS vector pUraRas has the mRas-prey order (Figure 2.1). This may

create an artificial membrane binding site at the fusion point between the mRas and the i1, i1-i2 or i1-GGG-i2 prey, which leads to membrane localisation of the mRas thus autoactivation. This could explain why the VLCYCLF-i1 and CYCLF-i1 baits did not show autoactivation. The difference between these two baits and i1 is the seven (VLCYCLF) or five (CYCLF) additional amino acids, which would break up the artificial membrane binding site that caused the autoactivation. Likewise, this could also explain why i2, which would form a completely unrelated fusion site, did not show autoactivation. One possible solution would be to move the prey protein to the front of the mRas to remove any potential binding site at the mRas-prey fusion point.

All GCR1 derived preys were used in the rRRS so that comparison can be made. However, since i1, i1-i2 or i1-GGG-i2 showed autoactivation, only VLCYCLF-i1, CYCLF-i1 and i2 can be used to draw conclusions on the rRRS experiments. As shown clearly in figure 7.8, there was a much higher level of yeast growth for VLCYCLF-i1 and CYCLF-i1 than for pUraRas, indicating that both VLCYCLF-i1 and CYCLF-i1 interacted with *zf*-DHHC1, and their interactions were unlikely to be caused by pUraRas interacting with *zf*-DHHC1. However, the level of yeast growth for i2 was similar to that of the pUraRas, indicating that there was either very weak interaction between i2 and *zf*-DHHC1, or no interaction. Assuming that there was an interaction for *zf*-DHHC1 with i2, judging from the yeast growth level alone, the interaction for *zf*-DHHC1 with VLCYCLF-i1 and CYCLF-i1 was much stronger than with i2, which is consistent with the RRS result. However, the differences in yeast growth seen may not be due entirely to the strength of the protein-protein interaction, but also to the amount of bait and preys in the cell. Because both the bait and prey vectors do not contain any epitope tag, we could not check the expression level of the bait and preys. In addition, there was yeast growth on the Gal-L-U+M media, on which only prey is expected to be expressed. This is especially true for days 5 and 6 (Figure 7.8). As stated in section 6.3.2, this could be expected as methionine can not completely suppress bait expression, and the low level of expressed bait protein binding to the prey would lead to cell proliferation.

It is worth to mention that the verification of the interaction between GCR1 and *zf*-DHHC1 by pull-down was not performed, because the expression of *zf*-DHHC1 in *E.coli* was hampered, possibly due to its association with the cell membrane. This problem could possibly be resolved by using the intracellular regions of *zf*-DHHC1, e.g. the DHHC domain without the TM region, rather than the full length protein, in the pull-down.

7.3.3 *zf*-DHHC1 might be involved in the palmitoylation of GCR1

For most DHHC PATs, the DHHC domain is located between TM2 and TM3, which is predicted to put it on the cytosolic face of the membrane along with the Nter that precedes TM1 and the Cter that lies after TM4 (Keller *et al*, 2004; Valdez-Taubas and Pelham, 2005; Mitchell *et al*., 2006). To examine whether *zf*-DHHC1 has this feature, its membrane topography was predicted using three computational programs (section 7.2.1). Interestingly, *zf*-DHHC1 was predicted to be a four-transmembrane protein with both its Nter and Cter in the cotyplasm. Its *zf*-DHHC domain, which is facing the cytoplasm, is also located between TM2 and TM3, extending into TM3 (Figure 7.2). Therefore, the membrane topography of *zf*-DHHC1 fits very well in the topography model of a typical DHHC PAT. Furthermore, the amino acid sequence of *zf*-DHHC1 was aligned with two well studied yeast PATs, Swf1 and Erf2p, and it was found that the greatest homology is around the regions of the DHHC domain. The conserved amino acids within the DHHC domain that are present in Swf1, Erf2p, and other DHHC PATs e.g. Akr1, HIP14, DHHC9 (Resh, 2006), are also present in *zf*-DHHC1 (Figure 7.2). These data strongly suggest that *zf*-DHHC1 may function as a PAT in *Arabidopsis*.

Some proteins are palmitoylated on cysteine residues in the N-terminal or C-terminal tail, or in both domains (reviewed in Qanbar and Bouvier, 2003). Many GPCRs are palmitoylated on cysteines in their C-terminal tail, such as rhodopsin (Ovchinnikov *et al*., 1988), β_2-adrenergic receptor (O'Dowd *et al*., 1998), and CCR5 (Percherancier *et al*., 2001). However, palmitoylation of

GPCRs are not limited to the C-terinal tail of this type of receptor. Mutation of all the cysteines in the C-terminal tail of the rat μ-opioid receptor failed to affect palmitate incorporation, indicating that the palmitoylation site(s) exist outside this receptor domain (Chen *et al.*, 1998; Qanbar and Bouvier, 2003). In fact, some proteins are palmiotylated within their intracellular regions. For instance, GODZ in mice has been shown to mediate palmitoylation of the intracellular regions of $GABA_A$ receptor, which is required for normal assembly and function of the GABAergic inhibitory synapse (Keller *et al.*, 2004; Fang *et al.*, 2006). Many others are palmitoylated on cysteines near the membrane/cytoplasm border (Ponimaskin and Schmidt, 1998). For example, three yeast SNAREs, Snc1, Tlg1 and Syn8 are palmitoylated on cysteine residues close to the cytoplasmic end of the TM domain (Valdez-Taubas and Pelham, 2005). It has been suggested that the presence of a cysteine residue near the membrane/cytoplasm border in the TM domain is be an excellent predictor of palmitoylation (Linder and Deschenes, 2007; Figure 7.10). The position of the two cysteines in GCR1, and the strong interaction between *zf*-DHHC1 and the baits containing these two cysteines, together with the possible function of *zf*-DHHC1, indicate that *zf*-DHHC1 is very likely involved in the palmitoylation of GCR1.

Palmitoylated integral membrane proteins	TM domain
influenza virus (subtype H7) HA protein	VILWFSFGASCFLLLAIAMGLVFIC*V
murine leukemia virus envelope protein	FTTLISTIMGPLIILLLILLFGPC*ILN
Semliki forest virus E1 protein	ISGGLGAFAIGAILVLVVVTC*IGL
Newcastle disease virus F protein	LTSTSALITYIVLTIISLVFGILSLVLAC*YLM
human immunodeficiency virus receptor CD4	PMALIVLGGVAGLLLFIGLGIFFC*V
Semliki forest virus F protein	AATVSAVVGMSLLALISIFASC*YMLVAA

Figure 7.10 Palmitoylation sites in integral membrane proteins (taken from Ponimaskin and Schmidt, 1998). * indicates palmitoylated cysteine residues in the TM domain.

The most commonly used experimental approach to prove that a protein is palmitoylated is to monitor the incorporation of radiolabel palmitate (^{3}H-palmitate) into the protein of interest (Resh 2006). For example, the palmitoylation of the V_2 vasopressin receptor (V2R) was detected by incubating transfected cells with ^{3}H-palmitates and immunoprecipitating the receptor with an antibody raised against a portion of V2R (Sadeghi *et al*., 1997). Substitution of CC341/342 for serines of V2R eliminated its palmitoylation, whereas replacement of a single amino acid, C341S or C342S, restored partial palmitoylation. Therefore, this approach can be applied not only to assess whether GCR1 is palmitoylated, but also to confirm whether the two cysteines in TM1 are the palmitoylation sites, by examining the palmitoylation of the equivalent cysteine to serine mutations of GCR1. The DHHC and TM domains of a PAT have been revealed to be critically important for interaction with its substrate (Lobo *et al*., 2002; Keller *et al*, 2004). The truncation of the putative N- and C-terminal cytoplasmic domains of GODZ did not interfere with its binding to the γ2 subunit of $GABA_A$ receptor. In contrast, more extensive deletions that disrupted the N and C-terminal TM domains or internal deletions that included the DHHC domain abolished the interaction with the γ2 subunit (Keller *et al*., 2004). Likewise, the domain(s) of *zf*-DHHC1 responsible for its interaction with GCR1 can be mapped by using its TM or DHHC domain truncations as the preys and GCR1 as the bait in the RRS.

When considering the roles of palmitoylation, it is useful to distinguish between constitutive palmitoylation, which refers to the initial protein palmitoylation occurring in the biosynthetic and transport pathway, and dynamic palmitoylation, which corresponds to the regulation of the palmitoylation state of the protein once it has reached its site(s) of action (Mitchell *et al*., 2006). In many cases, constitutive palmitoylation appears to play an important role in the expression of functional GPCR on the cell surface. For example, mutations that removed C341 and C342 reduced expression of V2R at the cell surface by 30%, while palmitoylation of either one of the two cysteines was sufficient to restore cell surface expression to wild-type levels (Schulein *et al*., 1996; Sadeghi *et al*.,

1997). Some non-palmitoylated GPCR mutants have been found to be retained intracellularly, primarily in the ER (Resh, 2006). While constitutive palmitoylation is important for efficient expression of functional receptor on the cell surface, dynamic palmitoylation is believed to protect integral membrane proteins from the cellular quality-control machinery (Linder and Deschenes, 2007). For instance, when the palmitoylation of Tlg1 is prevented by the deletion of the gene for Swf1 or through the mutation of the cysteines in Tlg1, Tlg1 is mis-localised to the vacuole, where it is subsequently degraded (Valdez-Taubas and Pelham, 2005). More examples are the A1 adenosine receptor, CCR5, and Rous sarcoma virus glycoprotein, which have all been shown to be palmitoylated close to a TM domain and to be rapidly degraded when palmitoylation is blocked by mutation (reviewed in Valdez-Taubas and Pelham, 2005). It is intriguing to speculate that GCR1 may be palmitoylated by *zf*-DHHC1 to achieve its cell surface expression, or to protect itself from degradation.

Palmitoylation has been shown to influence GPCRs downstream signalling. When palmitoylation is blocked, coupling to G-proteins is defective for some GPCRs, such as the β_2-adrenergic receptor (O'Dowd *et al*, 1989). Site-directed mutagenesis of the cysteine residue in the C-terminal sequence of the β_2-adrenergic receptor uncouples the receptor from its associated G-protein Gs (O'Dowd *et al*, 1989). However, this phenomenon is not observed for all GPCRs tested, like the dopamine D1 receptor (Jin *et al*., 1997) and human A_1 adenosine receptor (Gao *et al*., 1999). It would be interesting to discover whether palmitoylation is critical for the coupling of GCR1 to GPA1 in *Arabidopsis*.

Palmitoylation is distinguished from the other two lipid modifications, myristoylation and prenylation, by its reversibility. PATs are responsible for palmitoylation, whereas protein palmitoyl thioesterases (APTs) carry out depalmitoylation by removing the thioester-linked palmitate from the palmitoylated proteins (Linder and Deschenes, 2007). Cycles of palmitoylation

and depalmitoylation occur in a regulated manner for many proteins (Linder and Deschenes, 2004; Smotrys and Linder, 2004; Roth *et al*., 2006). GPCRs are known to be extensively phosphorylated by a number of kinases, including PKA and GRKs. For many GPCRs, phorphorylation are closely related to receptor desensitization and internalization (Qanbar and Bouvier, 2003). It has been established that palmitoylation is a molecular switch regulating the accessibility of phosphorylation sites involved in the desensitization or internalization of a receptor. Protein palmitoylation could occlude nearby phophorylation sites and protect them from phosphorylation, whereas depalmitoylation is suggested to promote GPCR phosphorylation, leading to internalization or desensitization (or both) and decreased signalling capacity (Resh, 2006). If *zf*-DHHC1 is truly a GCR1 PAT, further research detailing the effect of GCR1 palmitoylation by *zf*-DHHC1 on GCR1 internalization and desensitization would be of great interest.

Chapter 8 General discussion

Because plants are sessile organisms, their survival depends on the ability to sense and respond to changes in their environment. Plants have developed sensitive mechanisms to detect and react to environmental changes and the molecular basis of these is beginning to be understood. It appears that these signal transduction pathways form complex and interacting networks that integrate multiple aspects of cellular behaviour. Many proteins interact stably or transiently with neighbouring proteins and act as integrators of crosstalk in these signalling networks. On this basis, studying protein-protein interactions is fundamental to the understanding of how these signalling networks work.

Significant progress has been made in the past decades in developing methods to study protein-protein interactions. As for GIPs, many of them are identified by Co-IP or pull-down linked to proteomics analysis including 2-dimensional gel electrophoresis (2-DE) and MS. For example, Co-IP followed by MS analysis revealed the interaction of metabotropic glutamate receptor 5 (mGluR5) with known regulatory proteins, as well as novel proteins that constitute the mGluR5-signaling complex (Farr *et al*., 2004). By combining affinity chromatography, large-scale immunoprecipitation and MS, Husi *et al*. (2000) successfully characterized the N-methyl-D-aspartate receptor (NMDAR) multiprotein complex which comprise 77 proteins organised into receptor, adaptor, signalling, cytoskeletal and novel proteins. More recently, several new methods have been developed and successfully employed in protein-protein interaction analysis. A good example is the rapid affinity-capture of signalling proteins (GRASP) which allowed the identification of an interactor of Rho guanine Nucleotide Dissociation Inhibitor (RhoGDI) when coupled with MS (Berman *et al*., 2004). The advantage of this new method is that it detects protein-protein interactions present in the cells as they exist in their native tissue microenvironment (Berman *et al*., 2004).

Although the development of new technologies has resulted in a large number of reports on the identification of GIPs, in fact, most of the GIPs identified so far have been discovered using the yeast two-hybrid system (Bockaert *et al*., 2004b). Given its wide use and great success in isolating novel protein-protein interactions (Golemis and Adams, 2005), large-scale screening using the yeast two-hybrid system was considered as a good starting point for the identification of novel GCR1 interactors. The conventional GAL4 system was not used in the project because we did not consider that it would be ideal for analysing membrane proteins as explained in section 1.2.3. Indeed, the interaction between GCR1 and GPA1 was not observed using the GAL4 system (Humphrey and Botella, 2001; Pandey and Assmann, 2004). Therefore, the RRS which has been used as a good alternative to the GAL4 system for analysing protein-protein interactions of membrane proteins (Broder *et al*., 1998; Kohler and muller, 2003; Kruse *et al*., 2006) was used in this project for the isolation of novel GCR1 interacting proteins.

8.1 The use of the RRS as an initial screening approach to identify novel interactors for GCR1

Three baits, i1-i2, i3 and Cter, comprising the intracellular loops i1, i2, i3 and the C-terminal region of GCR1 were constructed for library screening using the RRS. Autoactivation screening revealed that only the i1-i2 bait was suitable for the RRS screening. In the initial screening using the i1-i2 bait, 774 clones were shown to require expression of both bait and prey. Twenty seven genes were identified that are represented at varying frequencies in 728 of the interactors. The i1-i2 bait along with three new baits i1, i2 and i1-GGG-i2 were used to reconfirm interactions with plasmids rescued from clones representing 14 of the 27 genes. The use of the three new baits allowed us to eliminate interactions that are based on the sequences at the boundary generated between the non-contiguous i1 and i2 sequences of the i1-i2 bait, and to trace whether each potential interactor binds to i1, i2, neither or both.

Extensive reconfirmation screening using the RRS revealed that three proteins, TRX3, TRX4 and *zf*-DHHC1, interacted with not only the original i1-i2 bait but also the three new baits, indicating that they interact with both i1 and i2 of GCR1. Furthermore, the interactions for these three proteins with GCR1 were supported by the rRRS and 6xHis-pull-down assays (TRX4 only so far), demonstrating that the RRS does appear to have found novel protein-protein interactions that are not seen before in the GCR1 signalling pathway. The other 11 potential interactors did not show interaction with the three new baits, but nearly all of them showed strong interaction with the original i1-i2 bait, implying that the original interactions were robust. The absence of their interaction with the new baits might be due to the elimination of potential binding site generated at the boundary of the i1 and i2 sequences of the i1-i2 bait, or might due to other unknown reasons. It is worthy to note that we could not recover the plasmids for certain interactors and thus were unable to verify their interaction with the four baits. It is possible that some of these unverified interactors do interact with GCR1.

Although the RRS has successfully identified novel interactors for GCR1, it also has a few drawbacks. For instance, autoactivation of certain baits in the RRS or preys in the rRRS reduced the productivity of the experiment to some extent. A wealth of information has shown that the i3 and Cter regions of GPCRs are the important regions that are involved in G-protein coupling. These regions were chosen as baits for use in the library screening using the RRS but unfortunately both showed autoactivation. This might be the reason why some interactors such as GPA1, which might have been identified using the i3 and Cter baits, were missed when using the i1-i2 bait. The evidence supporting this hypothesis comes from the work presented by Pandey and Assmann (2004). Using a series of truncated mutants of GCR1 with full length GPA1 in the split-ubiquitin two-hybrid assay, they demonstrated that the interaction between GCR1 and GPA1 requires the presence of an intact i3, a free Cter, and/or some key amino acids from the beginning of i2 to the beginning of i3 in GCR1. In addition, the only mark to detect the protein-

protein interaction in the RRS is the growth of yeast cells at the restrictive temperature (Causier and Davies, 2002). Because reversion of the temperature-sensitive mutation of the yeast stain does happen occasionally, the temperature sensitivity of the yeast cells needs to be checked before carrying out the screening. In the RRS, the expression of the bait is controlled by the pMet425 promoter which allows bait expression only in the absence of methionine, whereas the expression of prey is controlled by the Gal1 promoter that allows prey expression only in the absence of glucose. During the course of the experiment, it appeared that the pMet425 promoter could not completely suppress bait expression, so a low level of yeast growth was observed on the Gal-L-U-+4M medium, which reduced the sensitivity of the experiment. Despite these weaknesses, the work presented in this thesis successfully demonstrates that the RRS is overall a good system to use for large-scale, high throughput protein-protein interaction screens for integral membrane proteins.

8.2 The contribution of the GCR1-TRX3, GCR1-TRX4 and GCR1-*zf*-DHHC1 interactions to our understanding of GCR1

Mammalian GPCRs are known to interact with not only G-proteins but also many other proteins. These proteins modify receptor activities or participate in G-protein-independent signalling pathways (Sun *et al*., 2007). Although no G-protein subunits have been isolated as GCR1 interactors, the project did identify some proteins that might modify GCR1 activity or participate in GCR1 signalling pathways. Two of these proteins, TRX3 and TRX4, are thioredoxins. Thioredoxin contains two cysteines in the conserved active site which allows it to reduce disulfide bridges of target proteins. Therefore it can interact with a wide variety of proteins, including G-protein subunits (Wong *et al*., 2004; Alkhalfioui *et al*., 2007) and various receptors (Makino *et al*., 1999; Hansen *et al*., 2006). For instance, it plays a critical role in modulating TNF or steroid receptor-mediated signalling events via redox regulation of receptor activities (Makino *et al*., 1999; Hansen *et al*., 2006). The data presented in chapter 6

demonstrated that TRX3 and TRX4 could bind to i1 and i2 of GCR1 in the RRS, rRRS and pull-down assays. The fact that TRX3 and TRX4 are predicted to be in the cytoplasm (Gelhaye *et al*., 2004a; Meyer *et al*., 2006) is consistent with the location of these regions of GCR1 used as baits. Thioredoxin normally reduces the cysteine residues on target proteins. GCR1 does not contain any cysteines in the intracellular regions but has two cysteines in the VLCYCLF motif at the cytoplasmic end of TM1 near i1. These cysteines could potentially be exposed to the cytoplasm if GCR1 undergoes a conformational change that shifts the TM1 boundary, making them a potential target for thioredoxin. We therefore tested whether TRX3 and TRX4 interact with the extended i1 bait containing part of the TM1 domain with these two cysteines (VLCYCLF-i1 and CYCLF-i1). Out data showed that TRX3 and TRX4 may have a stronger interaction with VLCYCLF-i1 and CYCLF-i1 than with i1, indicating that they could bind to the cysteines in TM1 and reduce one or both of these cysteines. In order to examine whether these cysteines are reduced by TRX3 and TRX4, we made a series of active site mutants for TRX3 and TRX4 to test their interaction with the GCR1 derived baits. The mutation in active site appeared to abolish binding in the RRS although we need to confirm this is not due to a failure of prey expression. The data obtained so far lead us to tentatively suggest that TRX3 and TRX4 could reduce one or both of these cysteines.

The reducing ability enables thioredoxin to function as a powerful antioxidant to detoxify ROS, such as O_3 and H_2O_2. In plants, several lines of evidence have indicated that thioredoxin acts as a key regulator in oxidative stress response (Santos and Rey, 2006). For example, thioredoxin has been found to reduce ROS levels during seed desiccation and germination as well as nodule development, and many enzymes involved in ROS scavenging processes have been isolated as thioredoxin targets (Serrate and Cejudo, 2003; Lee *et al*., 2005; Santos and Rey, 2006). Yeast complementation results demonstrated that both TRX3 and TRX4 were able to confer H_2O_2 tolerance in yeast *trx1, trx2* double-mutant (Mouaheb *et al*., 1998; Brehelin *et*

al., 2000). Given that GCR1 interacts with both thioredoixn and G-protein α subunit, and that thioredoixn and G-protein β subunit interact with each other and are both involved in oxidative stress response, a speculative model is proposed (Figure 6.22) which suggests that thioredoxin may function as a key regulator that modifies GCR1 and G-protein activities in response to various oxidative stress conditions. It is known that in addition to being toxicants ROS can also act as essential signalling molecules to regulate many biological processes (Mittler *et al*, 2004; Misra *et al*., 2007), and that GPCR can mediate ROS-related signal transduction (Thannickal and Fanburg, 2000; Pelletier *et al*., 2003). Therefore, it is also possible that thioredoxin acts as a redox sensor to regulate GCR1-mediated ROS-related signalling pathways.

The rRRS data demonstrated that TRX3 and TRX4 interact with all parts of GCR1. This suggests that they may not be targeting GCR1 with their reductive capabilities and could instead be binding to provide other functions to GCR1. Accumulating evidence indicates that independent of its role as a reducing catalyst, thioredoxin can function as a chaperone. This introduces the possibility that TRX3 and TRX4 may modulate GCR1 activity through their chaperone activities. In fact, many GIPs have been shown to behave as chaperones and are involved in surface expression of the newly synthesized or recycled GPCR, or targeting misfolded GPCR to its degradation pathway (Brady and Limbird, 2002; Ulloa-Aguirre *et al*., 2004; Free *et al*., 2007). Therefore, TRX3 and TRX4 may act as chaperones to regulate GCR1 targeting or signalling.

No direct interaction between thioredoxin and GPCRs has been reported to date, although the list of proteins shown to interact with GPCRs includes thioredoxin-like proteins, e.g. protein kinase C-interacting cousin of thioredoxin (PICOT) (Becamel *et al*., 2002). The identification of TRX3 and TRX4 as GCR1 interacting proteins is the first time a direct interaction between thioredoxins and a GPCR has been observed. Apparently, much work is required to elucidate the interacting mechanism between thioredoxin

and GCR1, which will provide further insight into the role of TRX3 and TRX4 in GCR1 signalling.

A DHHC domain containing protein *zf*-DHHC1 is the third protein that was found to interact with both i1 and i2 of GCR1. Growing evidence suggests that proteins with the DHHC domain are PATs that palmitoylate target protein at cysteine residues (Gleason *et al*., 2006; Resh 2006). The result presented in chapter 7 showed that *zf*-DHHC1 has predicted transmembrane topography that is shared by known DHHC PATs – four TM domains with the DHHC domain located between TM2 and TM3 and extending into TM3 (Smotrys and Linder 2004; Keller *et al*., 2004; Mitchell *et al*., 2006), suggesting a likely role of *zf*-DHHC1 as an *Arabidopsis* PAT. In a continuation of the work reported for *zf*-DHHC1, Dr. Baoxiu Qi has isolated homozygous *zf*-DHHC1 mutant lines in *Arabidopsis*. The mutant plant has a striking phenotype: dwarf, produces fewer flowers than wild type plant, and is almost infertile. Interestingly, the mutant plant of another *Arabidopsis* DHHC protein Tip Growth Defect 1 (TIP1), which has confirmed PAT activity, has a phenotype that is very similar to *zf*-DHHC1: reduction in cell size, plant size, fertility and growth polarity (Ryan *et al*., 1998; Hemsley *et al*., 2005). The phenotypic similarity between the TIP1 and *zf*-DHHC1 mutant supports the idea that *zf*-DHHC1 is a PAT in *Arabidopsis*. Many integral membrane proteins are palmitoylated at the cysteine residues close to the cytoplasmic end of the TM domain, or in the intracellular regions (Ponimaskin and Schmidt, 1998; Qanbar and Bouvier, 2003; Keller *et al.*, 2004; Valdez-Taubas and Pelham, 2005). As stated earlier, GCR1 does not contain any cysteine residue in the intracellular regions but has two cysteines in the VLCYCLF motif at the cytoplasmic end of TM1, which might be potential palmitoylation sites. Interestingly, *zf*-DHHC1 demonstrated a stronger interaction with the VLCYCLF-i1 and CYCLF-i1 baits which contain the two cysteines in TM1, than with the i1 bait in the RRS, indicating a strong interaction between *zf*-DHHC1 and the motif containing the two cysteines. These data lead us to speculate that *zf*-DHHC1 may be involved in the palmitoylation of GCR1. In

mammals, nearly all GPCRs are palmitoylated with functional implications to receptor signalling (Tobin and Wheatley, 2004). For instance, some GPCRs require palmitoylation to achieve cell surface expression, to couple to downstream effectors, or to protect themselves from degradation (O'Dowd *et al*, 1989; Resh, 2006; Linder and Deschenes, 2007). Whether this is also the case for GCR1 still awaits exploration. It has been suggested that palmitoylation is essential for normal plant cell growth (Hemsley *et al*., 2005). It is intriguing to speculate that the powerful effect of palmitoylation on plant cell growth is achieved through the regulation of the activity of GCR1 and other substrates by DHHC PATs, including *zf*-DHHC1 and TIP1.

Prior to this study, no protein other than GPA1 has been identified to directly interact with GCR1. The most significant finding of this project is the three novel interactors TRX3, TRX4 and *zf*-DHHC1 that interacts with at least the i1 an i2 regions of GCR1. This finding expands our knowledge of the "GCR1 interactome", contributes to the growing understanding of the GCR1 signalling network, and opens many new avenues for experiment.

8.3 Future directions

In this thesis, we provide direct evidence for the interaction of GCR1 with three proteins TRX3, TRX4 and *zf*-DHHC1, both in yeast and *in vitro*. Since these are plant proteins, their interactions are expected to occur *in vivo* in plants, which can be tested by Co-IP (section 1.2.2). Plant expression constructs for GCR1, TRX3, TRX4 and *zf*-DHHC1 have been made, and transient expression of TRX3 and TRX4 has been detected in tobacco (supplementary chapter). Future work will be required to g*et al*l of these proteins expressed in *Arabidopsis* and to test their interactions by Co-IP.

Defining the domains or amino acids of the interactors and GCR1 that are responsible or required for their interactions would be an important experiment to conduct as the result would help elucidate the interacting

mechanism. This can be performed by using a series of truncated or point mutations in the RRS, rRRS or pull-down assays. Since the RRS prey vector and the rRRS bait and prey vectors do not contain any epitope tag, it is worthwhile to add epitope tags to these vectors so that the expression of the bait and prey can be easily monitored and quantified. In addition, a number of experiments can be carried out to test the hypotheses regarding the involvement of TRX3, TRX4 and *zf*-DHHC1 in GCR1 signalling, for instance, whether TRX3, TRX4 and GCR1 are involved in oxidative stress response, and whether GCR1 can be palmitoylated by *zf*-DHHC1 and if so how palmitoylation affects GCR1 activities.

The time constrains of this project did not allow detailed investigation of the reason for i3 and Cter autoactivation in the RRS. Future work can be conducted to resolve the autoactivation issue so that both can be used as baits in the RRS screens for GCR1 interactors. Given the importance of the i3 and Cter regions of mammalian GPCRs in their interactions with downstream effectors, the use of these two baits in the RRS screens are expected to identify more GCR1 interactors.

In addition to protein-protein interaction analysis, many other approaches can be employed to study GCR1, such as resolving its 3D structure or carrying out genetics studies. New insights gained will help us to decipher its signalling pathway and bring its physiological importance to light.

References

Abe M., Fujiwara M., Kurotani K.-i., Yokoi S., and Shimamoto K. (2008). Identification of dynamin as an interactor of Rice GIGANTEA by tandem affinity purification (TAP). Plant Cell Physiol. **49(3),** 420-432.

Adams M. D., Celniker S. E., Holt R. A., Evans C. A., Gocayne J. D., Amanatides P. G., Scherer S. E., Li P. W., Hoskins R. A., and Galle R. F. (2000). The genome sequence of Drosophila melanogaster. Science **287(5461),** 2185-2195.

Alkhalfioui F., Renard M., Vensel W. H., Wong J., Tanaka C. K., Hurkman W. J., Buchanan B. B., and Montrichard F. (2007). Thioredoxin-linked proteins are reduced during germination of Medicago truncatula seeds. Plant Physiol. **144(3),** 1559-1579.

Alves I. D., Park C. K., and Hruby V. J. (2005). Plasmon resonance methods in GPCR signaling and other membrane events. Curr Protein Pept Sci. **6 (4),** 293–312.

Andoh T., Chock P. B., and Chiueh C. C. (2002). The roles of thioredoxin in protection against oxidative stress-induced apoptosis in SH-SY5Y Cells. J. Biol. Chem. **277(12),** 9655-9660.

Andreeva A. V., and Kutuzov M. A. (2001). Do plants have rhodopsin after all? A mystery of plant G protein-coupled signalling. Plant Physiology and Biochemistry **39(12),** 1027-1035.

Apone F., Alyeshmerni N., Wiens K., Chalmers D., Chrispeels M. J., and Colucci G. (2003). The G-protein-coupled receptor GCR1 regulates DNA synthesis through activation of phosphatidylinositol-specific phospholipase C. Plant Physiol. **133(2),** 571-579.

Arifuzzaman M., Maeda M., Itoh A., Nishikata K., Takita C., Saito R., Ara T., Nakahigashi K., Huang H.-C., and Hirai A. (2006). Large-scale identification of protein-protein interaction of Escherichia coli K-12. Genome Res. **16(5),** 686-691.

Arner E. S. J., and Holmgren A. (2000). Physiological functions of thioredoxin and thioredoxin reductase. European Journal of Biochemistry **267(20),** 6102-6109.

Arredondo-Peter R., Hargrove M. S., Moran J. F., Sarath G., and Klucas R. V. (1998). Plant hemoglobins. Plant Physiol **118(4),** 1121-1125.

Arredondo-Peter R., Hargrove M. S., Moran J. F., Sarath G., and Klucas R. V. (1998). Plant hemoglobins. Plant Physiol **118(4),** 1121-1125.

Assmann S. M. (2002). Heterotrimeric and unconventional GTP binding proteins in plant cell signaling. Plant Cell **14(90001),** S355-373.

Assmann S. M. (2005). G Proteins Go Green: A Plant G Protein Signaling FAQ Sheet. Science **310(5745),** 71-73.

Aswath C. R., and Kim S. H. (2005). Another story of MADS-box genes – their potential in plant biotechnology. Plant Growth Regulation **46(2),** 177-188.

Azarkan M., Huet J., Baeyens-Volant D., Looze Y., and Vandenbussche G. (2007). Affinity chromatography: A useful tool in proteomics studies. Journal of Chromatography B **849(1-2),** 81-90.

Ballicora M. A., Frueauf J. B., Fu Y., Schurmann P., and Preiss J. (2000). Activation of the potato tuber ADP-glucose pyrophosphorylase by thioredoxin. J. Biol. Chem. **275(2),** 1315-1320.

Balmer Y., and Schurmann P. (2001). Heterodimer formation between thioredoxin f and fructose 1,6-bisphosphatase from spinach chloroplasts. FEBS Letters **492(1-2)**, 58-61.

Bancos S., Nomura T., Sato T., Molnar G., Bishop G. J., Koncz C., Yokota T., Nagy F., and Szekeres M. (2002). Regulation of transcript levels of the Arabidopsis cytochrome P450 genes involved in brassinosteroid biosynthesis. Plant Physiol. **130(1),** 504-513.

Bargmann C. I. (1998). Neurobiology of the Caenorhabditis elegans genome. Science **282(5396),** 2028-2033.

Becamel C., Alonso G., Galeotti N., Demey E., Jouin P., Ullmer C., Dumuis A., Bockaert J., and Marin P. (2002). Synaptic multiprotein complexes associated with 5-HT_{2C} receptors: a proteomic approach. EMBO J. **21(10),** 2332–2342.

Berman D. M., Shih I.-M., Burke L.-A., Veenstra T. D., T Y. Z., Conrads h. P., Kwon S. W., Hoang V., Yu L.-R., Zhou M., Kurman R. J., Petricoin E. F., and Liotta L. A. (2004). Profiling the activity of G proteins in patient-derived tissues by rapid affinity-capture of signal transduction proteins (GRASP). PROTEOMICS **4(3),** 812-818.

Bernard V., Decossas M., Liste I., and Bloch B. (2006). Intraneuronal trafficking of G-protein-coupled receptors in vivo. Trends in Neurosciences **29(3),** 140-147.

Bhattacharya M., Babwah A. V., and Ferguson S. S. G. (2004). Small GTP-binding protein-coupled receptors. Biochem. Soc. Trans. **32(Pt 6),** 1040-1044.

Bockaert J., and Pin J. P. (1999). Molecular tinkering of G protein-coupled receptors: an evolutionary success. EMBO J. **18(7),** 1723-1729.

Bockaert J., Fagni L., Dumuis A., and Marin P. (2004a). GPCR interacting proteins (GIP). Pharmacology & therapeutics **103(3),** 203-221.

Bockaert J., Marin P., Dumuis A., and Fagni L. (2003). The 'magic tail' of G protein-coupled receptors: an anchorage for functional protein networks. FEBS Letters **546(1),** 65-72.

Bockaert J., Roussignol G., Bacamel C., Gavarini S., Joubert L., Dumuis A., Fagni L., and Marin P. (2004b). GPCR-interacting proteins (GIPs): nature and functions. Biochem. Soc. Trans. **32(5),** 851-855.

Boisvert F.-M., Cote J., Boulanger M.-C., and Richard S. (2003). A proteomic analysis of arginine-methylated protein complexes. Mol Cell Proteomics **2(12),** 1319-1330.

Booker F. L., Burkey K. O., Overmyer K., and Jones A. M. (2004). Differential responses of G-protein Arabidopsis thaliana mutants to ozone. New Phytologist **162(3),** 633-641.

Bower M. S., Matias D. D., Fernandes-Carvalho E., Mazzurco M., Gu T., Rothstein S. J., and Goring D. R. (1996). Two members of the thioredoxin-h family interact with the kinase domain of a Brassica S locus receptor kinase. Plant Cell **8(9),** 1641-1650.

Brady A. E., and Limbird L. E. (2002). G protein-coupled receptor interacting proteins: Emerging roles in localization and signal transduction. Cellular Signalling **14(4),** 297-309.

Brehelin C., Laloi C., Setterdahl A., Knaff D., and Meyer Y. (2004). Cytosolic, mitochondrial thioredoxins and thioredoxin reductases in Arabidopsis thaliana. Photosynthesis Research **79,** 295-304.

Brehelin C., Mouaheb N., Verdoucq L., Lancelin J. M., and Meyer Y. (2000). Characterization of determinants for the specificity of Arabidopsis thioredoxins h in yeast complementation. J Biol Chem **275(41),** 31641-31647.

Broder Y. C., Katz S., and Aronheim A. (1998). The Ras recruitment system, a novel approach to the study of protein-protein interactions. Current Biology **8(20),** 1121-1130.

Broin M., Cuine S., Eymery F., and Rey P. (2002). The plastidic 2-cysteine peroxiredoxin is a target for a thioredoxin involved in the protection of the photosynthetic apparatus against oxidative damage. Plant Cell **14(6),** 1417-1432.

Brown-Augsburger P., Hartshorn K., Chang D., Rust K., Fliszar C., Welgus H. G., and Crouch E. C. (1996). Site-directed mutagenesis of Cys-15 and Cys-20 of pulmonary surfactant protein D. Expression of a trimeric protein with altered anti-viral properties. J Biol Chem **271(23),** 13724-13730.

Bulenger S., Marullo S., and Bouvier M. (2005). Emerging role of homo- and heterodimerization in G-protein-coupled receptor biosynthesis and maturation. Trends in Pharmacological Sciences **26(3),** 131-137.

Carmel-Harel O., and Storz G. (2000). Roles of the glutathione- and thioredoxin-dependent reduction systems in the Escherichia coli and saccharomyces cerevisiae responses to oxidative stress. Annual Review of Microbiology **54(1),** 439-461.

Cashmore A. R., Jarillo J. A., Wu Y.-J., and Liu D. (1999). Cryptochromes: blue light receptors for plants and animals. Science **284(5415),** 760-765.

Causier B. (2004). Studying the interactome with the yeast two-hybrid system and mass spectrometry. Mass Spectrometry Reviews **23(5),** 350-367.

Causier B., and Davies B. (2002). Analysing protein-protein interactions with the yeast two-hybrid system. Plant Molecular Biology **50,** 855-870.

Chang I.-F. (2006). Mass spectrometry-based proteomic analysis of the epitope-tag affinity purified protein complexes in eukaryotes. PROTEOMICS **6(23),** 6158-6166.

Charest P. G., and Bouvier M. (2003). Palmitoylation of the V2 vasopressin receptor carboxyl tail enhances β-Arrestin recruitment leading to efficient receptor endocytosis and ERK1/2 activation. J. Biol. Chem. **278(42),** 41541-41551.

Chen C. P., and Rost B. (2002). Long membrane helices and short loops predicted less accurately. Protein Sci **11(12),** 2766-2773.

Chen C., Shahabi V., Xu W., and Liu-Chen L. Y. (1998). Palmitoylation of the rat μ opioid receptor. FEBS Letters **441,** 148-152.

Chen J.-G., and Jones A. M. (2004). AtRGS1 function in Arabidopsis thaliana. Methods in Enzymology **Volume 389,** 338-350.

Chen J.-G., Gao Y., and Jones A. M. (2006a). Differential roles of Arabidopsis heterotrimeric G-protein subunits in modulating cell division in roots. Plant Physiol. **141(3),** 887-897.

Chen J.-G., Pandey S., Huang J., Alonso J. M., Ecker J. R., Assmann S. M., and Jones A. M. (2004). GCR1 can act independently of heterotrimeric G-protein in response to brassinosteroids and gibberellins in Arabidopsis seed germination. Plant Physiol. **135(2),** 907-915.

Chen J.-G., Willard F. S., Huang J., Liang J., Chasse S. A., Jones A. M., and Siderovski D. P. (2003). A seven-transmembrane RGS protein that modulates plant cell proliferation. Science **301(5640),** 1728-1731.

Chen Y., Ji F., and Liang J. (2006b). Overexpression of the regulator of G-protein signalling protein enhances ABA-mediated inhibition of root elongation and drought tolerance in Arabidopsis. Journal of Experimental Botany **57,** 2101-2110.

Chen Y., Ji F., Xie H., Liang J., and Zhang J. (2006c). The regulator of G-protein signaling proteins involved in sugar and abscisic acid signaling in Arabidopsis seed germination. Plant Physiol. **140(1),** 302-310.

Cherezov V., Rosenbaum D. M., Hanson M. A., Rasmussen S. G. F., Thian F. S., Kobilka T. S., Choi H.-J., Kuhn P., Weis W. I., Kobilka B. K., and Stevens R. C. (2007). High-resolution crystal structure of an engineered human β2-adrenergic G protein coupled receptor. Science **318(5854),** 1258-1265.

Chicchi G. G., Graziano M. P., Koch G., Hey P., Sullivan K., Vicario P. P., and Cascieri M. A. (1997). Alterations in receptor activation and divalent cation activation of agonist binding by deletion of intracellular domains of the glucagon receptor. J. Biol. Chem. **272(12),** 7765-7769.

Cho M.-H., Corea O. R. A., Yang H., Bedgar D. L., Laskar D. D., Anterola A. M., Moog-Anterola F. A., Hood R. L., Kohalmi S. E., Bernards M. A., Kang C., Davin L. B., and Lewis N. G. (2007). Phenylalanine biosynthesis in Arabidopsis thaliana. J. Biol. Chem. **282(42),** 30827-30835.

Cho W. (2001). Membrane Targeting by C1 and C2 Domains. Journal of Biological Chemistry. **276(35),** 32407-32410.

Christie J. M. (2007). Phototropin blue-light receptors. Annual Review of Plant Biology **58(1),** 21-45.

Coates P., and Hall P. (2003). The yeast two-hybrid system for identifying protein-protein interactions. The Journal of Pathology **199,** 4-7.

Colley N. J., Baker E. K., Stamnes M. A., and Zuker C. S. (1991). The cyclophilin homolog ninaA is required in the secretory pathway. Cell **67(2),** 255-263.

Colucci G., Apone F., Alyeshmerni N., Chalmers D., and Chrispeels M. J. (2002). GCR1, the putative Arabidopsis G protein-coupled receptor gene is cell cycle-regulated, and its overexpression abolishes seed dormancy and shortens time to flowering. Proceedings of the National Academy of Sciences **99(7),** 4736-4741.

Cook J. V. F., and Eidne K. A. (1997). An intramolecular disulfide bond between conserved extracellular cysteines in the gonadotropin-releasing hormone receptor is essential for binding and activation. Endocrinology **138(7),** 2800-2806.

Cotecchia S., Fanelli F., and Costa T. (2003). Constitutively active G protein-coupled receptor mutants:Implications on receptor function and drug action. ASSAY and Drug Development Technologies **1(2),** 311-316.

Creighton T. E. (2003). Proteins: Structures and Molecular Properties. (W.H. Freeman and Company).

Crespo P., and Leon J. (2000). Ras proteins in the control of the cell cycle and cell differentiation. Cellular and Molecular Life Sciences **57(11),** 1613-1636.

Criekinge W. V., and Beyaert R. (1999). Yeast two-hybrid: state of the art. Biological Procedures Online **2(1),** 1-38.

Damdimopoulos A. E., Miranda-Vizuete A., Pelto-Huikko M., Gustafsson J.-A., and Spyrou G. (2002). Human mitochondrial thioredoxin. Involvement in mitochondrial membrane potential and cell death. J. Biol. Chem. **277(36),** 33249-33257.

Davies M. N., Secker A., Freitas A. A., Mendao M., Timmis J., and Flower D. R. (2007). On the hierarchical classification of G protein-coupled receptors. Bioinformatics **23(23),** 3113-3118

Devi L. A. (2005). The G protein-coupled receptors handbook. (Humana Press).

Devoto A., Hartmann H. A., Piffanelli P., Elliott C., Simmons C., Taramino G., Goh C.-S., Cohen F. E., Emerson B. C., Schulze-Lefert P., and Panstruga R. (2003). Molecular phylogeny and evolution of the plant-specific seven-transmembrane MLO family. Journal of Molecular Evolution **56(1),** 77-88.

Devoto A., Piffanelli P., Nilsson I., Wallin E., Panstruga R., von Heijne G., and Schulze-Lefert P. (1999). Topology, subcellular localization, and sequence diversity of the Mlo family in plants. J. Biol. Chem. **274(49),** 34993-35004.

Dievart A., and Clark S. E. (2004). LRR-containing receptors regulating plant development and defense. Development **131(2),** 251-261.

Dortay H., Mehnert N., Burkle L., Schmulling T., and Heyl A. (2006). Analysis of protein interactions within the cytokinin-signaling pathway of Arabidopsis thaliana. FEBS Journal **273(20),** 4631-4644.

Earley K. W., Haag J. R., Pontes O., Opper K., Juehne T., Song K., and Pikaard C. S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. The Plant Journal 45(4), 616-629.

Einarson M. B., Pugacheva E. N., and Orlinick J. R. (2007). GST Pull-down. Cold Spring Harbor Protocols **2007(16),** pdb.prot4757.

Ellis C. (2004). The state of GPCR research in 2004. Nat Rev Drug Discov **3(7),** 577-626.

Elofsson A., and Heijne G. v. (2007). Membrane Protein Structure: Prediction versus Reality. Annual Review of Biochemistry **76(1),** 125-140.

Englbrecht C. C., Schoof H. and Bohm S. (2004). Conservation, diversification and expansion of C_2H_2 zinc finger proteins in the Arabidopsis thaliana genome. BMC Genomics. **5,** 39.

Escriba P. V., Wedegaertner P. B., Goni F. M., and Vogler O. (2007). Lipid-protein interactions in GPCR-associated signaling. Biochimica et Biophysica Acta (BBA) - Biomembranes **1768(4),** 836-852.

Fang C., Deng L., Keller C. A., Fukata M., Fukata Y., Chen G., and Luscher B. (2006). GODZ-mediated palmitoylation of $GABA_A$ receptors is required for normal assembly and function of GABAergic inhibitory synapses. J. Neurosci. **26(49),** 12758-12768.

Fang L., Stevens J. L., Berk A. J., and Spindler K. R. (2004). Requirement of Sur2 for efficient replication of mouse adenovirus type 1. Journal of Virology **78(23),** 12888-12900.

Farr C. D., Gafken P. R., Norbeck A. D., Doneanu C. E., Stapels M. D., Barofsky D. F., Minami M., and Saugstad J. A. (2004). Proteomic analysis of native metabotropic glutamate receptor 5 protein complexes reveals novel molecular constituents. Journal of Neurochemistry **91(2),** 438-450.

Felberbaum-Corti M., Morel E., Cavalli V., Vilbois F., and Gruenberg J. (2007). The redox sensor TXNL1 plays a regulatory role in fluid phase endocytosis. PLoS ONE **2(11),** e1144.

Ferguson S. S. G. (2001). Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling. Pharmacol Rev **53(1),** 1-24.

Fields S. (2001). Two-hybrid and related systems. In: Encyclopedia of life sciences. John Wiley and Sons, Ltd: Chichester. [DOI: 10.1038/npg.els.0000981].

Fields S., and Song O.-k. (1989). A novel genetic system to detect protein–protein interactions. Nature **340(6230),** 245-246.

Finkelstein R. R., Gampala S. S. L., and Rock C. D. (2002). Abscisic acid signaling in seeds and seedlings. Plant Cell **14(90001),** S15-45.

Frankel P., Aronheim A., Kavanagh E., Balda M. S., Matter K., Bunney T. D., and Marshall C. J. (2005). RalA interacts with ZONAB in a cell density-dependent manner and regulates its transcriptional activity. The EMBO Journal **24(1),** 54-62.

Fredriksson R., and Schioth H. B. (2005). The repertoire of G-protein-coupled receptors in fully sequenced genomes. Mol Pharmacol **67(5),** 1414-1425.

Fredriksson R., Lagerstrom M. C., Lundin L.-G., and Schioth H. B. (2003). The G-protein-coupled receptors in the human genome form five main families.

Phylogenetic analysis, paralogon Groups, and fingerprints. Mol Pharmacol **63(6),** 1256-1272.

Free R. B., Hazelwood L. A., Cabrera D. M., Spalding H. N., Namkung Y., Rankin M. L., and Sibley D. R. (2007). D1 and D2 dopamine receptor expression is regulated by direct interaction with the chaperone protein calnexin. J. Biol. Chem. **282(29),** 21285-21300.

Fukushimaa Y., Saitohb T., Anaia M., Ogiharaa T., Inukaia K., Funakia M., Sakodaa H., Onishia Y., Onoa H., Fujishiroa M., Ishikawaa T., Takatac K., Nagaia R., Omataa M., and Asanoa T. (2001). Palmitoylation of the canine histamine H2 receptor occurs at Cys305 and is important for cell surface targeting. Biochimica et Biophysica Acta (BBA)/Molecular Cell Research **1539,** 181-191.

Gainetdinov R. R., Premont R. T., Bohn L. M., Lefkowitz R. J., and Caron M. G. (2004). Desensitization of G protein-coupled receptors and neuronal functions. Annual Review of Neuroscience **27(1),** 107-144.

Gao Z., Ni Y., Szabo G., and Linden J. (1999). Palmitoylation of the recombinant human A1 adenosine receptor: enhanced proteolysis of palmitoylation-deficient mutant receptors. Biochem. J. **342(2),** 387-395.

Gelhaye E., Rouhier N., and Jacquot J. P. (2004a). The thioredoxin h system of higher plants. Plant Physiol Biochem **42(4),** 265-271.

Gelhaye E., Rouhier N., Gerard J., Jolivet Y., Gualberto J., Navrot N., Ohlsson P.-I., Wingsle G., Hirasawa M., Knaff D. B., Wang H., Dizengremel P., Meyer Y., and Jacquot J.-P. (2004b). A specific form of thioredoxin h occurs in plant mitochondria and regulates the alternative oxidase. Proceedings of the National Academy of Sciences **101(40),** 14545-14550.

Gelhaye E., Rouhier N., Navrot N., and Jacquot J. P. (2005). The plant thioredoxin system. Cellular and Molecular Life Sciences **62(1),** 24-35.

Gelhaye E., Rouhier N., Vlamis-Gardikas A., Girardet J. M., Sautiere P. E., Sayzet M., Martin F., and Jacquot J. P. (2003). Identification and characterization of a third thioredoxin h in poplar. Plant Physiology and Biochemistry **41,** 629-635.

George S. R., O'Dowd B. F., and Lee S. P. (2002). G-Protein-coupled receptor oligomerization and its potential for drug discovery. Nat Rev Drug Discov **1(10),** 808-820.

Gleason E. J., Lindsey W. C., Kroft T. L., Singson A. W., and L'Hernault S. W. (2006). spe-10 Encodes a DHHC-CRD Zinc-Finger Membrane Protein Required for Endoplasmic Reticulum/Golgi Membrane Morphogenesis During Caenorhabditis elegans Spermatogenesis. Genetics **172(1),** 145-158.

Golemis E. A., and Adams P. D. (2005). Protein-protein interactions. (Cold Spring Harbor Laboratory Press).

Graaf C. d., Foata N., Engkvist O., and Rognan D. (2008). Molecular modeling of the second extracellular loop of G-protein coupled receptors and its implication on structure-based virtual screening. Proteins: Structure, Function, and Bioinformatics **71(2),** 599-620.

Gurevich E., and Gurevich V. (2006). Arrestins: ubiquitous regulators of cellular signaling pathways. Genome Biology **7(9),** 236.

Hall A. (2000). GTPases. (Oxford University Press).

Hammer M. M., Wehrman T. S., and Blau H. M. (2007). A novel enzyme complementation-based assay for monitoring G-protein-coupled receptor internalization. The Journa of The Federation of American Societies for Experimental Biology (FASEB J.) **21(14),** 3827-3834.

Hansen J. M., Zhang H., and Jones D. P. (2006). Mitochondrial thioredoxin-2 has a key role in determining tumor necrosis factor-α-induced reactive oxygen species generation, NF-*κ*B activation, and apoptosis. Toxicol. Sci. **91(2),** 643-650.

Harding P., Hadingham T., McDonnell J., and Watts A. (2006). Direct analysis of a GPCR-agonist interaction by surface plasmon resonance. European Biophysics Journal **35,** 709-712.

Hauger R. L., Smith R. D., Braun S., Dautzenberg F. M., and Catt K. J. (2000). Rapid agonist-induced phosphorylation of the human CRF receptor, type 1: A potential mechanism for homologous desensitization. Biochemical and Biophysical Research Communications **268,** 572-576.

He R., Browning D. D., and Ye R. D. (2001). Differential roles of the NPXXY motif in formyl peptide receptor signaling. The Journal of Immunology **166(6),** 4099-4105.

Hemsley P. A., Kemp A. C., and Grierson C. S. (2005). The TIP GROWTH DEFECTIVE1 S-acyl transferase regulates plant cell growth in Arabidopsis. Plant Cell **17(9),** 2554 - 2563.

Hennemann H., Vassen L., Geisen C., Eilers M., and Moroy T. (2003). Identification of a novel KRAB box domain protein - KRIM-1 - that interacts with c-Myc and inhibits its oncogenic activity. J. Biol. Chem. M207196200.

Hochuli E., Dobeli H., and Schacher A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. Journal of Chromatography A **411,** 177-184.

Holmgren A. (1995). Thioredoxin structure and mechanism: conformational changes on oxidation of the active-site sulfhydryls to a disulfide. Structure **3(3),** 239-243.

Hooley R. (1998). Plant hormone perception and action: a role for G-protein signal transduction? Philosophical Transactions of the Royal Society B: Biological Sciences **353(1374),** 1425-1430.

Horn F., Weare J., Beukers M. W., Horsch S., Bairoch A., Chen W., Edvardsen O., Campagne F., and Vriend G. (1998). GPCRDB: an information system for G protein-coupled receptors. Nucl. Acids Res. **26(1),** 275-279.

Huber H. E., Russel M., Model P., and Richardson C. C. (1986). Interaction of mutant thioredoxins of Escherichia coli with the gene 5 protein of phage T7. The redox capacity of thioredoxin is not required for stimulation of DNA polymerase activity. J. Biol. Chem. **261(32),** 15006-15012.

Hubsman M., Yudkovsky G., and Aronheim A. (2001). A novel approach for the identification of protein-protein interaction with integral membrane proteins. Nucl. Acids Res. **29(4),** e18.

Humphrey T. V., and Botella J. R. (2001). Re-evaluation of the cytokinin receptor role of the Arabidopsis gene GCR1. Journal of Plant Physiology **158,** 645-653.

Husi H., Ward M. A., Choudhary J. S., Blackstock W. P., and Grant S. G. N. (2000). Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. Nat Neurosci **3(7),** 661-669.

Initiative T. A. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature **408(6814),** 796-815.

Insel P. A., Tang C.-M., Hahntow I., and Michel M. C. (2007). Impact of GPCRs in clinical medicine: Monogenic diseases, genetic variants and drug targets. Biochimica et Biophysica Acta (BBA) - Biomembranes **1768(4),** 994-1005.

Iwasaki Y., Fujisawa Y., and Kato H. (2003). Function of heterotrimeric G protein in gibberellin signaling. Journal of Plant Growth Regulation **22(2),** 126-133.

Jee C., Vanoaica L., Lee J., Park B. J., and Ahnn J. (2005). Thioredoxin is related to life span regulation and oxidative stress response in Caenorhabditis elegans. Genes to Cells **10(12),** 1203-1210.

Jeng M.-F., Campbell A. P., Begley T., Holmgren A., Case D. A., Wright P. E., and Dyson H. J. (1994). High-resolution solution structures of oxidized and reduced Escherichia coli thioredoxin. Structure **2(9),** 853-868.

Jensen R. B., Lykke-Andersen K., Frandsen G. I., Nielsen H. B., Haseloff J., Jespersen H. M., Mundy J., and Skriver K. (2000). Promiscuous and specific phospholipid binding by domains in ZAC, a membrane-associated Arabidopsis protein with an ARF GAP zinc finger and a C2 domain. Plant Mol Biol **44(6),** 799-814.

Jin H., Zastawny R., George S. R., and O'Dowd B. F. (1997). Elimination of palmitoylation sites in the human dopamine D1 receptor does not affect receptor-G protein interaction. European Journal of Pharmacology **324(1),** 109-116.

Johnston C. A., Taylor J. P., Gao Y., Kimple A. J., Grigston J. C., Chen J.-G., Siderovski D. P., Jones A. M., and Willard F. S. (2007b). GTPase acceleration as the rate-limiting step in Arabidopsis G protein-coupled sugar signaling. Proceedings of the National Academy of Sciences **104(44),** 17317-17322.

Johnston C. A., Temple B. R., Chen J.-G., Gao Y., Moriyama E. N., Jones A. M., Siderovski D. P., and Willard F. S. (2007a). Comment on "A G protein coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid". Science **318(5852),** 914c.

Jones A. M. (2002). G-protein-coupled signaling in Arabidopsis. Current Opinion in Plant Biology **5(5),** 402-407.

Jones A. M., and Assmann S. M. (2004). Plants: the latest model system for G-protein research. EMBO reports **5(6),** 572–578.

Joo J. H., Wang S., Chen J. G., Jones A. M., and Fedoroff N. V. (2005). Different signaling and cell death roles of heterotrimeric G protein α and β subunits in the Arabidopsis oxidative stress response to ozone. Plant Cell **17(3),** 957–970.

Josefsson L.-G. (1999). Evidence for kinship between diverse G-protein coupled receptors. Gene **239(2),** 333-340.

Josefsson L.-G., and Rask L. (1997). Cloning of a putative G-protein-coupled receptor from Arabidopsis thaliana. European Journal of Biochemistry **249(2),** 415-420.

Jurado P., de Lorenzo V., and Fernández L. A. (2006). Thioredoxin fusions increase folding of single chain Fv antibodies in the cytoplasm of Escherichia coli: Evidence that chaperone activity is the prime effect of thioredoxin. Journal of Molecular Biology **357(1),** 49-61.

Kahsay R. Y., Gao G., and Liao L. (2005). An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. Bioinformatics **21(9),** 1853-1858.

Keller C. A., Yuan X., Panzanelli P., Martin M. L., Alldred M., Sassoe-Pognetto M., and Luscher B. (2004). The γ2 Subunit of GABAA Receptors Is a Substrate for Palmitoylation by GODZ. J. Neurosci. **24(26),** 5881-5891.

Kern R., Malki A., Holmgren A., and Richarme G. (2003). Chaperone properties of Escherichia coli thioredoxin and thioredoxin reductase. Biochem. J. **371(3),** 965-972.

Kim M. C., Lee S. H., Kim J. K., Chun H. J., Choi M. S., Chung W. S., Moon B. C., Kang C. H., Park C. Y., Yoo J. H., Kang Y. H., Koo S. C., Koo Y. D., Jung J. C., Kim S. T., Schulze-Lefert P., Lee S. Y., and Cho M. J. (2002b). Mlo, a modulator of plant defense and cell death, is a novel calmodulin-binding protein. J. Biol. Chem. **277(22),** 19304-19314.

Kim M. C., Panstruga R., Elliott C., Muller J., Devoto A., Yoon H. W., Park H. C., Cho M. J., and Schulze-Lefert P. (2002a). Calmodulin interacts with MLO protein to regulate defence against mildew in barley. Nature **416(6879),** 447-451.

Kleanthous C. (2000). Protein-protein recognition. (Oxford University Press).

Kohler F., and Muller K. M. (2003). Adaptation of the Ras-recruitment system to the analysis of interactions between membrane-associated proteins. Nucleic Acids Research **31,** e28.

Kruse C., Hanke S., Vasiliev S., and Hennemann H. (2006). Protein-protein interaction screening with the Ras-recruitment system. Signal Transduction **6(3),** 198-208.

Kuceraa B., Cohna M. A., and Leubner-Metzgera G. (2005). Plant hormone interactions during seed dormancy release and germination. Seed Science Research **15,** 281-307

Kumar J. K., Tabor S., and Richardson C. C. (2004). Proteomic analysis of thioredoxin-targeted proteins in Escherichia coli. Proceedings of the National Academy of Sciences **101(11),** 3759-3764.

Kundu S., Trent J. T., 3rd, and Hargrove M. S. (2003). Plants, humans and hemoglobins. Trends Plant Sci **8(8),** 387-393.

Kundu S., Trent J. T., 3rd, and Hargrove M. S. (2003). Plants, humans and hemoglobins. Trends Plant Sci **8(8),** 387-393.

Lacal J. C. (1997). Regulation of proliferation and apoptosis by Ras and Rho GTPases through specific phospholipid-dependent signaling. FEBS Letters **410(1),** 73-77.

Lanctot P. M., Leclerc P. C., Clement M., Auger-Messier M., Escher E., Leduc R., and Guillemette G. (2005). Importance of N-glycosylation positioning for cell-surface expression, targeting, affinity and quality control of the human AT1 receptor. Biochem J. **15(Pt1),** 367–376.

Lapik Y. R., and Kaufman L. S. (2003). The Arabidopsis cupin domain protein AtPirin1 interacts with the G protein α-subunit GPA1 and regulates seed germination and early seedling development. Plant Cell **15(7),** 1578-1590.

Lee J., Sayegh J., Daniel J., Clarke S., and Bedford M. T. (2005a). PRMT8, a new membrane-bound tissue-specific member of the protein arginine methyltransferase family. J. Biol. Chem. **280(38),** 32890-32896.

Lee M.-Y., Shin K.-H., Kim Y.-K., Suh J.-Y., Gu Y.-Y., Kim M.-R., Hur Y.-S., Son O., Kim J.-S., and Song E. (2005b). Induction of thioredoxin is required for nodule development to reduce reactive oxygen species levels in soybean roots. Plant Physiol. **139(4),** 1881-1889.

Lehner B., Semple J. I., Brown S. E., Counsell D., Campbell R. D., and Sanderson C. M. (2004). Analysis of a high-throughput yeast two-hybrid system and its use to predict the function of intracellular proteins encoded within the human MHC class III region. Genomics **83(1),** 153-167.

Li B., Cong F., Tan C. P., Wang S. X., and Goff S. P. (2002). Aph2, a protein with a zf-DHHC motif, interacts with c-Abl and has pro-apoptotic activity. J. Biol. Chem. **277(32),** 28870-28876.

Linder M. E., and Deschenes R. J. (2004). Model organisms lead the way to protein palmitoyltransferases. Journal of Cell Science **117(4),** 521-526.

Linder M. E., and Deschenes R. J. (2007). Palmitoylation: policing protein stability and traffic. Nat Rev Mol Cell Biol **8(1),** 74-84.

Liu F., Ni W., Griffith M. E., Huang Z., Chang C., Peng W., Ma H., and Xie D. (2004). The ASK1 and ASK2 genes are essential for Arabidopsis early development. Plant Cell **16(1),** 5-20.

Liu F., Ni W., Griffith M. E., Huang Z., Chang C., Peng W., Ma H., and Xie D. (2004). The ASK1 and ASK2 genes are essential for Arabidopsis early development. Plant Cell **16(1),** 5-20.

Liu H., Nishitoh H., Ichijo H., and Kyriakis J. M. (2000). Activation of apoptosis signal-regulating kinase 1 (ASK1) by tumor necrosis factor receptor-associated factor 2 requires prior dissociation of the ASK1 inhibitor thioredoxin. Mol. Cell. Biol. **20(6),** 2198-2208.

Liu X., Yue Y., Li B., Nie Y., Li W., Wu W.-H., and Ma L. (2007b). A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. Science **315(5819),** 1712-1716.

Liu X., Yue Y., Li W., and Ma L. (2007a). Response to comment on "A G protein coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid". Science **318(5852),** 914d.

Liu Y., and Min W. (2002). Thioredoxin promotes ASK1 ubiquitination and degradation to inhibit ASK1-mediated apoptosis in a redox activity-independent manner. Circ Res **90(12),** 1259-1266.

Lobo S., Greentree W. K., Linder M. E., and Deschenes R. J. (2002). Identification of a Ras palmitoyltransferase in Saccharomyces cerevisiae. J. Biol. Chem. **277(43),** 41268-41273.

Lopper M. E., and Keck J. L. (2007). Protein–protein interactions: identification. In: Encyclopedia of life sciences. John Wiley and Sons, Ltd: Chichester. [DOI: 10.1002/9780470015902.a9780470020491].

Lundstrom K., Wagner R., Reinhart C., Desmyter A., Cherouati N., Magnin T., Zeder-Lutz G., Courtot M., Prual C., André N., Hassaine G., Michel H., Cambillau C., and Pattus F. (2006). Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems. Journal of Structural and Functional Genomics **7(2),** 77-91.

Ma H. (1994). GTP-binding proteins in plants: new members of an old family. Plant Molecular Biology **26(5),** 1611-1636.

Makino Y., Yoshikawa N., Okamoto K., Hirota K., Yodoi J., Makino I., and Tanaka H. (1999). Direct association with thioredoxin allows redox regulation of glucocorticoid receptor function. J. Biol. Chem. **274(5),** 3182-3188.

Mantovani R. (1998). A survey of 178 NF-Y binding CCAAT boxes. Nucl. Acids Res. **26(5),** 1135-1143.

Marinissen M. J., and Gutkind J. S. (2001). G-protein-coupled receptors and signaling networks: emerging paradigms. Trends in Pharmacological Sciences **22(7),** 368-376.

McCoy J. G., Arabshahi A., Bitto E., Bingman C. A., Ruzicka F. J., Frey P. A., and Phillips G. N. (2006). Structure and mechanism of an ADP-glucose phosphorylase from Arabidopsis thaliana. Biochemistry **45(10),** 3154-3162.

McLatchie L. M., Fraser N. J., Main M. J., Wise A., Brown J., Thompson N., Solari R., Lee M. G., and Foord S. M. (1998). RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor. Nature **393(6683),** 333-339.

Meyer Y., Riondet C., Constans L., Abdelgawwad M., Reichheld J., and Vignols F. (2006). Evolution of redoxin genes in the green lineage. Photosynthesis Research **89,** 179-192.

Miller W. E., Houtz D. A., Nelson C. D., Kolattukudy P. E., and Lefkowitz R. J. (2003). G-protein-coupled Receptor (GPCR) kinase phosphorylation and β-arrestin recruitment regulate the constitutive signaling activity of the human cytomegalovirus US28 GPCR. J. Biol. Chem. **278(24),** 21663-21671.

Milligan G., and White J. H. (2001). Protein-protein interactions at G-protein-coupled receptors. Trends in Pharmacological Sciences **22(10),** 513-518.

Misra S., Wu Y., Venkataraman G., Sopory S. K., and Tuteja N. (2007). Heterotrimeric G-protein complex and G-protein-coupled receptor from a legume (Pisum sativum): role in salinity and heat stress and cross-talk with phospholipase C. The Plant Journal **51(4),** 656-669.

Mitchell D. A., Vasudevan A., Linder M. E., and Deschenes R. J. (2006). Thematic review series: lipid posttranslational modifications. protein palmitoylation by a family of DHHC protein S-acyltransferases. Journal of Lipid Research **47(6),** 1118-1127.

Mitsui A., Hamuro J., Nakamura H., Kondo N., Hirabayashi Y., Ishizaki-Koizumi S., Hirakawa T., Inoue T., and Yodoi J. (2002). Overexpression of human thioredoxin in transgenic mice controls oxidative stress and life span. Antioxidants & Redox Signaling **4(4),** 693-696. .

Mittler R., Vanderauwera S., Gollery M., and Van Breusegem F. (2004). Reactive oxygen gene network of plants. Trends in Plant Science **9(10),** 490-498.

Mizutani M., Ward E., and Ohta D. (1998). cytochrome p450 superfamily in Arabidopsis thaliana: isolation of cDNAs, Differential Expression, and RFLP mapping of multiple cytochromes P450. Plant Molecular Biology **37,** 39-52.

Mora-Garcia S., Rodriguez-Suarez R., and Wolosiuk R. A. (1998). Role of electrostatic interactions on the affinity of thioredoxin for target proteins. J. Biol. Chem. **273(26),** 16273-16280.

Moriyama E., Strope P., Opiyo S., Chen Z., and Jones A. (2006). Mining the Arabidopsis thaliana genome for highly-divergent seven transmembrane receptors. Genome Biology **7(10),** R96.

Mouaheb N., Thomas D., Verdoucq L., Monfort P., and Meyer Y. (1998). In vivo functional discrimination between plant thioredoxins by heterologous expression in the yeast Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences **95(6),** 3312-3317.

Mukherjee S., Bal S., and Saha P. (2001). Protein interaction maps using yeast two-hybrid assay. Current Science **81(5),** 458-464.

Nakagawa M., Miyamoto T., Kusakabe R., Takasaki S., Takao T., Shichida Y., and Tsuda M. (2001). O-Glycosylation of G-protein-coupled receptor, octopus rhodopsin: Direct analysis by FAB mass spectrometry. FEBS Letters **496(1),** 19-24.

Nambi P., and Aiyar N. (2003). G protein-coupled receptors in drug discovery. ASSAY and Drug Development Technologies **1(2),** 305-310.

Nawy T., Malamy J. E., Thongrod S., Jung J., and Benfey P. N. (2002). The MADS box gene AGL42 is expressed in the quiescent center and maintains root meristem organization. In 13th International Conference on Arabidopsis Research.

Neer E. J. (1995). Heterotrimeric G proteins: Organizers of transmembrane signals. Cell **80(2),** 249-257.

Nemoto W., and Toh H. (2003). Prediction of interfaces for GPCR oligomer. Genome Informatics **13,** 512-513.

Neves S. R., Ram P. T., and Iyengar R. (2002). G protein pathways. Science **296(5573),** 1636-1639.

Nishiyama A., Ohno T., Iwata S., Matsui M., Hirota K., Masutani H., Nakamura H., and Yodoi J. (1999). Demonstration of the interaction of thioredoxin with p40phox, a phagocyte oxidase component, using a yeast two-hybrid system. Immunology Letters **68(1),** 155-159.

O'Connor C. D., and Hames B. D. (2008). Proteomics. (Scion Publishing Limited).

O'Dowd B. F., Hnatowich M., Caron M. G., Lefkowitz R. J., and Bouvier M. (1989). Palmitoylation of the human β_2-adrenergic receptor. Mutation of Cys341 in the carboxyl tail leads to an uncoupled nonpalmitoylated form of the receptor. J. Biol. Chem. **264(13),** 7564-7569.

Okuno Y., Yang J., Taneishi K., Yabuuchi H., and Tsujimoto G. (2006). GLIDA: GPCR-ligand database for chemical genomic drug discovery. Nucl. Acids Res. **34(suppl_1),** D673-677.

Omerovic J., Laude A., and Prior I. (2007). Ras proteins: paradigms for compartmentalised and isoform-specific signalling. Cellular and Molecular Life Sciences **64(19),** 2575-2589.

Ovchinnikov Y. A., Abdulaev N. G., and Bogachuk. A. S. (1988). Two adjacent cysteine residues in the C-terminal cytoplasmic fragment of bovine rhodopsin are palmitylated. FEBS Letters **230(1-2),** 1-5.

Palczewski K., Kumasaka T., Hori T., Behnke C. A., Motoshima H., Fox B. A., Trong I. L., Teller D. C., Okada T., Stenkamp R. E., Yamamoto M., and Miyano M. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. Science **289(5480),** 739-745.

Pandey S., and Assmann S. M. (2004). The Arabidopsis putative G protein-coupled receptor GCR1 interacts with the G protein α subunit GPA1 and regulates abscisic acid signaling. Plant Cell **16(6),** 1616-1632.

Pandey S., Chen J.-G., Jones A. M., and Assmann S. M. (2006). G-Protein complex mutants are hypersensitive to abscisic acid regulation of germination and postgermination development. Plant Physiol. **141(1),** 243-256.

Pandit S. B., and Srinivasan N. (2003). Survey for g-proteins in the prokaryotic genomes: Prediction of functional roles based on classification. Proteins: Structure, Function, and Genetics **52(4),** 585-597.

Pao C. S., and Benovic J. L. (2005). Structure/function analysis of $_{\alpha 2A}$-adrenergic receptor interaction with G protein-coupled receptor kinase 2. J. Biol. Chem. **280(12),** 11052-11058.

Parker M. S., Wong Y. Y., and Parker S. L. (2008). An ion-responsive motif in the second transmembrane segment of rhodopsin-like receptors. Amino Acids**. [**DOI 10.1007/s00726-00008-00637-00726].

Pedrajas J. R., Kosmidou E., Miranda-Vizuete A., Gustafsson J.-A., Wright A. P. H., and Spyrou G. (1999). Identification and functional characterization of a novel mitochondrial thioredoxin system in Saccharomyces cerevisiae. J. Biol. Chem. **274(10),** 6366-6373.

Perazzolli M., Dominici P., Romero-Puertas M. C., Zago E., Zeier J., Sonoda M., Lamb C., and Delledonne M. (2004). Arabidopsis nonsymbiotic hemoglobin AHb1 modulates nitric oxide bioactivity. Plant Cell **16(10),** 2785-2794.

Perazzolli M., Dominici P., Romero-Puertas M. C., Zago E., Zeier J., Sonoda M., Lamb C., and Delledonne M. (2004). Arabidopsis nonsymbiotic hemoglobin AHb1 modulates nitric oxide bioactivity. Plant Cell **16(10),** 2785-2794.

Percherancier Y., Planchenault T., Valenzuela-Fernandez A., Virelizier J.-L., Arenzana-Seisdedos F., and Bachelerie F. (2001). Palmitoylation-dependent control of degradation, life span, and membrane expression of the CCR5 Receptor. J. Biol. Chem. **276(34),** 31936-31944.

Perfus-Barbeoch L., Jones A. M., and Assmann S. M. (2004). Plant heterotrimeric G protein function: insights from Arabidopsis and rice mutants. Current Opinion in Plant Biology **7(6),** 719-731.

Peterson F. C., Lytle B. L., Sampath S., Vinarov D., Tyler E., Shahan M., Markley J. L., and Volkman B. F. (2005). Solution structure of thioredoxin h1 from Arabidopsis thaliana. Protein Sci **14(8),** 2195-2200.

Phizicky E. M., and Fields S. (1995). Protein-protein interactions: methods for detection and analysis. Microbiology and Molecular Biology Reviews **59(1),** 94-123.

Picard D. (1999). Nuclear Receptors: A Practical Approach. (Oxford University Press).

Plakidou-Dymock S., Dymock D., and Hooley R. (1998). A higher plant seven-transmembrane receptor that influences sensitivity to cytokinins. Current Biology **8(6),** 315-324.

Polek T. C., Talpaz M., and Spivak-Kroizman T. (2006). The TNF receptor, RELT, binds SPAK and uses it to mediate p38 and JNK activation. Biochemical and Biophysical Research Communications **343(1),** 125-134.

Politis E. G., Roth A. F., and Davis N. G. (2005). Transmembrane topology of the protein palmitoyl transferase Akr1. J Biol Chem **280(11),** 10156 - 10163.

Ponimaskin E., and Schmidt M. F. G. (1998). Domain-structure of cytoplasmic border region is main determinant for palmitoylation of influenza virus hemagglutinin (H7). Virology **249(2),** 325-335.

Potter R. M., Maestas D. C., Cimino D. F., and Prossnitz E. R. (2006). Regulation of N-formyl peptide receptor signaling and trafficking by individual carboxyl-terminal serine and threonine residues. J Immunol **176(9),** 5418-5425.

Prather P. L. (2004). Inverse agonists: tools to reveal ligand-specific conformations of G protein-coupled receptors. Sci. STKE **2004(215),** pe1.

Puig O., Caspary F., Rigaut G., Rutz B., Bouveret E., Bragado-Nilsson E., Wilm M., and Seraphin B. (2001). The tandem affinity purification (TAP) method: A general procedure of protein complex purification. Methods **24(3),** 218-229.

Qian B., Soyer O. S., Neubig R. R., and Goldstein R. A. (2003). Depicting a protein's two faces: GPCR classification by phylogenetic tree-based HMMs. FEBS Letters **554, 1-2),** 95-99.

Qin J., Clore G. M., and Gronenborn A. M. (1994). The high-resolution three-dimensional solution structures of the oxidized and reduced states of human thioredoxin. Structure **2(6),** 503-522.

Ramachandran S., and Cerione R. A. (2006). How GPCRs hit the switch. Nat Struct Mol Biol **13(9),** 756-757.

Rasmussen S. G. F., Choi H.-J., Rosenbaum D. M., Kobilka T. S., Thian F. S., Edwards P. C., Burghammer M., Ratnala V. R. P., Sanishvili R., Fischetti R. F., Schertler G. F. X., Weis W. I., and Kobilka B. K. (2007). Crystal structure of the human β2 adrenergic G-protein-coupled receptor. Nature **450(7168),** 383-387.

Ray K., Ghosh S. P., and Northup J. K. (2004). The role of cysteines and charged amino acids in extracellular loops of the human Ca2+ receptor in cell surface expression and receptor activation processes. Endocrinology **145(8),** 3892-3903.

Rebollo A., and Martinez-A C. (1999). Ras proteins: recent advances and new functions. Blood **94(9),** 2971-2980.

Referencing electronic sources. Pierce: Proteomics pathways – protein interactions. (2005). [Online]. http://www.piercenet.com/Proteomics/

Referencing electronic sources. Robertus J.D., home page, section of CH 339K, Biochemistry. (2008). [Online]. http://courses.cm.utexas.edu/jrobertus/ch339k/overheads-1/ch5_affinity.jpg

Referencing electronic sources. TAIR - The Arabidopsis Information Resource. (2005). [Online]. http://www.arabidopsis.org

Referencing electronic sources. TMbase - A database of membrane spanning protein segments. (2008). [Online]. http://www.ch.embnet.org/software/tmbase/TMBASE_doc.html

Reichheld J.-P., Meyer E., Khafif M., Bonnard G., and Meyer Y. (2005). AtNTRB is the major mitochondrial thioredoxin reductase in Arabidopsis thaliana. FEBS Letters **579(2),** 337-342.

Reinheckel T., Sitte N., Ullrich O., Kuckelkorn U., Davies K. J., and Grune T. (1998). Comparative resistance of the 20S and 26S proteasome to oxidative stress. Biochem. J. **335(3),** 637-642.

Resh M. D. (2006). Palmitoylation of ligands, receptors, and intracellular signaling molecules. Sci STKE **2006(359),** re14.

Ritz D., Patel H., Doan B., Zheng M., Aslund F., Storz G., and Beckwith J. (2000). Thioredoxin 2 is involved in the oxidative stress response in Escherichia coli. J Biol Chem **275(4),** 2505-2512.

Rivera-Madrid R., Mestres D., Marinho P., Jacquot J., Decottignies P., Miginiac-Maslow M., and Meyer Y. (1995). Evidence for five divergent thioredoxin h sequences in Arabidopsis thaliana. Proceedings of the National Academy of Sciences **92(12),** 5620-5624.

Rizo J. and Sudhof T.C. (1998). C_2-domains, Structure and Function of a Universal Ca^2-binding Domain. Journal of Biological Chemistry, **273(26),** 15879-15882).

Rock C. D. (2000). Pathways to abscisic acid-regulated gene expression. New Phytologist **148,** 357-396.

Rohila J. S., Chen M., Cerny R., and Fromm M. E. (2004). Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. The Plant Journal **38(1),** 172-181.

Rohila J. S., Chen M., Chen S., Chen J., Cerny R., Dardick C., Canlas P., Xu X., Gribskov M., Kanrar S., Zhu J.-K., Ronald P., and Fromm M. E. (2006).

Proteinprotein interactions of tandem affinity purification-tagged protein kinases in rice. The Plant Journal **46,** 1-13.

Rosenbaum D. M., Cherezov V., Hanson M. A., Rasmussen S. G. F., Thian F. S., Kobilka T. S., Choi H.-J., Yao X.-J., Weis W. I., Stevens R. C., and Kobilka B. K. (2007). GPCR engineering yields high-resolution structural insights into β2-adrenergic receptor function. Science **318(5854),** 1266-1273.

Roth A. F., Feng Y., Chen L., and Davis N. G. (2002). The yeast DHHC cysteine-rich domain protein Akr1p is a palmitoyl transferase. J. Cell Biol. **159(1),** 23-28.

Roth A. F., Wan J., Bailey A. O., Sun B., Kuchar J. A., Green W. N., Phinney B. S., Yates J. R., and Davis N. G. (2006). Global analysis of protein palmitoylation in yeast. Cell **125(5),** 1003 - 1013.

Rouhier N., and Jacquot J.-P. (2003). Molecular and catalytic properties of a peroxiredoxin-glutaredoxin hybrid from Neisseria meningitidis. FEBS Letters **554(1-2),** 149-153.

Rubio V., Shen Y., Saijo Y., Liu Y., Gusmaroli G., Dinesh-Kumar S. P., and Deng X. W. (2005). An alternative tandem affinity purification strategy applied to Arabidopsis protein complex isolation. The Plant Journal **41(5),** 767-778.

Ryan E., Grierson C. S., Cavell A., Steer M., and Dolan L. (1998). TIP1 is required for both tip growth and non-tip growth in Arabidopsis. New Phytologist **138(1),** 49-58.

Sadeghi H. M., Innamorati G., Dagarag M., and Birnbaumer M. (1997). Palmitoylation of the V2 vasopressin receptor. Mol Pharmacol **52(1),** 21-29.

Saitoh M., Nishitoh H., Fujii M., Takeda K., Tobiume K., Sawada Y., Kawabata M., Miyazono K., and Ichijo H. (1998). Mammalian thioredoxin is a direct inhibitor of apoptosis signal-regulating kinase (ASK) 1. EMBO J **17(9),** 2596–2606.

Santos C. V. D., and Rey P. (2006). Plant thioredoxins are key actors in the oxidative stress response. Trends in Plant Science **11(7),** 329-334.

Schenk H., Klein M., Erdbrugger W., Droge W., and Schulze-Osthoff K. (1994). Distinct effects of thioredoxin and antioxidants on the activation of transcription factors NF-*κ*B and AP-1. Proceedings of the National Academy of Sciences **91(5),** 1672-1676.

Schoneberg T., Schulz A., Biebermann H., Hermsdorf T., Rompler H., and Sangkuhl K. (2004). Mutant G-protein-coupled receptors as a cause of human diseases. Pharmacology & Therapeutics **104(3),** 173-206.

Schulein R., Liebenhoff U., Muller H., Birnbaumer M., and Rosenthal W. (1996). Properties of the human arginine vasopressin V2 receptor after site-directed mutagenesis of its putative palmitoylation site. Biochem. J. **313(2),** 611-616.

Schurmann P., and Jacquot J. P. (2000). Plant thioredoxin system revisited. Annual Review of Plant Physiology and Plant Molecular Biology **51(1),** 371-400.

Schwacke R., Schneider A., van der Graaff E., Fischer K., Catoni E., Desimone M., Frommer W. B., Flugge U.-I., and Kunze R. (2003). ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. Plant Physiol. **131(1),** 16-26.

Seethala R., and Fernandes P. B. (2001). Handbook of Drug Screening. (Informa Healthcare).

Seitz H., Hutschenreiter S., Hultschig C., Zeilinger C., Zimmermann B., Kleinjung F., Schuchhardt J., Eickhoff H., and Herberg F. W. (2006). Differential binding studies applying functional protein microarrays and surface plasmon resonance. PROTEOMICS **6(19),** 5132-5139.

Shiu S.-H., and Bleecker A. B. (2001). Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. Proceedings of the National Academy of Sciences **98(19),** 10763-10768.

Silvio A. d., Imbriano C., and Mantovani R. (1999). Dissection of the NF-Y transcriptional activation potential. Nucl. Acids Res. **27(13),** 2578-2584.

Simon M. I., Strathmann M. P., and Gautam N. (1991). Diversity of G proteins in signal transduction. Science **252(5007),** 802-808.

Slepak V. Z. (2000). Application of surface plasmon resonance for analysis of protein-protein interactions in the G protein-mediated signal transduction pathway. Journal of Molecular Recognition **13(1),** 20-26.

Smotrys J. E., and Linder M. E. (2004). Palmitoylation of intracellular signaling proteins: regulation and function. Annual Review of Biochemistry **73(1),** 559-587.

Sonnhammer E. L. L., Heijne G. v., and Krogh A. (1998). A hidden markov model for predicting transmembrane helices in protein sequences. In Proceedings of the 6th International Conference on Intelligent Systems for Molecular Biology. (AAAI Press).

Spyrou G., Enmark E., Miranda-Vizuete A., and Gustafsson J. (1997). Cloning and expression of a novel mammalian thioredoxin. J Biol. Chem. **272(5),** 2936-2941.

Spyrou G., Enmark E., Miranda-Vizuete A., and Gustafsson J. (1997). Cloning and expression of a novel mammalian thioredoxin. J Biol Chem **272(5),** 2936-2941.

Stagljar I., and Fields S. (2002). Analysis of membrane protein interactions using yeast-based technologies. Trends in Biochemical Sciences **27(11),** 559-563.

Sun Y., McGarrigle D., and Huang X.-Y. (2007). When a G protein-coupled receptor does not couple to a G protein. Molecular BioSystems **3(12),** 849-854.

Svensson M. J., and Larsson J. (2007). Thioredoxin-2 affects lifespan and oxidative stress in Drosophila. Hereditas **144(1),** 25-32.

Tabor S., Huber H. E., and Richardson C. C. (1987). Escherichia coli thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. J. Biol. Chem. **262(33),** 16212-16223.

Takai Y., Sasaki T., and Matozaki T. (2001). Small GTP-binding proteins. Physiological Reviews **81(1),** 153-208.

Takemaru K.-I., Yamaguchi S., Lee Y. S., Zhang Y., Carthew R. W., and Moon R. T. (2003). Chibby, a nuclear β-catenin-associated antagonist of the Wnt/Wingless pathway. Nature **422(6934),** 905-909.

Tang J., Gary J. D., Clarke S., and Herschman H. R. (1998). PRMT 3, a type I protein arginine N-methyltransferase that differs from PRMT1 in its oligomerization, subcellular localization, substrate specificity, and regulation. J. Biol. Chem. **273(27),** 16935-16945.

Taylor P. D., Toseland C. P., Attwood T. K., and Flower D. R. (2006). Beta barrel trans-membrane proteins: Enhanced prediction using a Bayesian approach. Bioinformation **1(6),** 231–233.

Tcherpakov M., Bronfman F. C., Conticello S. G., Vaskovsky A., Levy Z., Niinobe M., Yoshikawa K., Arenas E., and Fainzilber M. (2002). The p75 neurotrophin receptor interacts with multiple MAGE proteins. J. Biol. Chem. **277(51),** 49101-49104.

Teller D. C., Okada T., Behnke C. A., Palczewski K., and Stenkamp R. E. (2001). Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-Protein-Coupled Receptors (GPCRs). Biochemistry **40(26),** 7761–7772.

Temple B. R. S., and Jones A. M. (2007). The plant heterotrimeric G-protein complex. Annual Review of Plant Biology **58(1),** 249-266.

Terrillon S., and Bouvier M. (2004). Roles of G-protein-coupled receptor dimerization EMBO reports **5(1),** 30–34.

Thannickal V. J., and Fanburg B. L. (2000). Reactive oxygen species in cell signaling. Am J Physiol Lung Cell Mol Physiol **279(6),** L1005-1028.

Thomas W. G., and Qian H. (2003). Arresting angiotensin type 1 receptors. Trends in Endocrinology and Metabolism **14,** 130-136.

Torrecilla I., Spragg E. J., Poulin B., McWilliams P. J., Mistry S. C., Blaukat A., and Tobin A. B. (2007). Phosphorylation and regulation of a G protein-coupled receptor by protein kinase CK2. J. Cell Biol. **177(1),** 127-137.

Trotter E. W., and Grant C. M. (2005). Overlapping roles of the cytoplasmic and mitochondrial redox regulatory systems in the yeast Saccharomyces cerevisiae. Eukaryotic Cell **4(2),** 392-400.

Tusnady G. E., and Simon I. (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. Journal of Molecular Biology **283(2),** 489-506.

Ullah H., Chen J.-G., Young J. C., Im K.-H., Sussman M. R., and Jones A. M. (2001). Modulation of cell proliferation by heterotrimeric G protein in Arabidopsis. Science **292(5524),** 2066-2069.

Ulloa-Aguirre A., Janovick J. A., Leanos-Miranda A., and Conn P. M. (2004). Misrouted cell surface GnRH receptors as a disease aetiology for congenital isolated hypogonadotrophic hypogonadism. Human Reproduction Update **10(2),** 177-192.

Valdez-Taubas J., and Pelham H. (2005). Swf1-dependent palmitoylation of the SNARE Tlg1 prevents its ubiquitination and degradation. The EMBO Journal **24,** 2524–2532.

Vanderbeld B., and Kelly G. M. (2000). New thoughts on the role of the βγ subunit in G protein signal transduction. Biochem. Cell Biol. **78(5),** 537–550.

Verdoucq L., Vignols F., Jacquot J.-P., Chartier Y., and Meyer Y. (1999). In vivo characterization of a thioredoxin h target protein defines a new peroxiredoxin family. J. Biol. Chem. **274(28),** 19714-19722.

Verhagen A. M., Ekert P. G., Pakusch M., Silke J., Connolly L. M., Reid G. E., Moritz R. L., Simpson R. J., and Vaux D. L. (2000). Identification of DIABLO, a mammalian protein that promotes apoptosis by binding to and antagonizing IAP proteins. Cell **102(1),** 43-53.

Vidalain P.-O., Boxem M., Ge H., Li S., and Vidal M. (2004). Increasing specificity in high-throughput yeast two-hybrid experiments. Methods **32(4),** 363-370.

Vitale A. (2002). Physical methods. Plant Molecular Biology **50(6),** 825-836.

von Heijne G. (2006). Membrane-protein topology. Nat Rev Mol Cell Biol **7(12),** 909-918.

Walhout A. J. M., Boulton S. J., and Vidal M. (2000). Yeast two-hybrid systems and protein interaction mapping projects for yeast and worm. Yeast **17(2),** 88-94.

Walsh M. T., Foley J. F., and Kinsella B. T. (1998). Characterization of the role of N-linked glycosylation on the cell signaling and expression of the human thromboxane A2 receptor alpha and beta isoforms. The Journal of Pharmacology and Experimental Therapeutics **286(2),** 1026-1036.

Wangensteen O. S., Chueca A., Hirasawa M., Sahrawy M., Knaff D. B., and Lopez Gorge J. (2001). Binding features of chloroplast fructose-1,6-bisphosphatase-thioredoxin interaction. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology **1547(1),** 156-166.

Warpeha K. M., Lateef S. S., Lapik Y., Anderson M., Lee B.-S., and Kaufman L. S. (2006). G-protein-coupled receptor 1, G-protein Gα-subunit 1, and prephenate dehydratase 1 are required for blue light-induced production of phenylalanine in etiolated Arabidopsis. Plant Physiol. **140(3),** 844-855.

Warpeha K. M., Upadhyay S., Yeh J., Adamiak J., Hawkins S. I., Lapik Y. R., Anderson M. B., and Kaufman L. S. (2007). The GCR1, GPA1, PRN1, NF-Y signal chain mediates both blue light and abscisic acid responses in Arabidopsis. Plant Physiol. **143(4),** 1590-1600.

Watson S., and Arkinstall S. (1994). The G-protein linked receptor factsbook. (Academic Press).

Wess J. (1998). Molecular basis of receptor/G-protein-coupling selectivity. Pharmacology & Therapeutics **80(3),** 231-264.

Wieland T., and Michel M. C. (2005). Can a GDP-liganded G-protein be active? Mol Pharmacol **68(3),** 559-562.

Wollman E. E., d'Auriol L., Rimsky L., Shaw A., Jacquot J. P., Wingfield P., Graber P., Dessarps F., Robin P., and Galibert F. (1988). Cloning and expression of a cDNA for human thioredoxin. J. Biol. Chem. **263(30),** 15506-15512.

Wong J. H., Cai N., Balmer Y., Tanaka C. K., Vensel W. H., Hurkman W. J., and Buchanan B. B. (2004). Thioredoxin targets of developing wheat seeds identified by complementary proteomic approaches. Phytochemistry **65(11),** 1629-1640.

Wong S. K. F. (2003). G protein selectivity is regulated by multiple intracellular regions of GPCRs. Neurosignals **12(1),** 1-12.

Yamanaka H., Maehira F., Oshiro M., Asato T., Yanagawa Y., Takei H., and Nakashima Y. (2000). A possible interaction of thioredoxin with VDUP1 in HeLa cells detected in a yeast two-hybrid system. Biochem Biophys Res Commun **271(3),** 796-800.

Yoshioka J., Schreiter E. R., and Lee R. T. (2006). Role of thioredoxin in cell growth through interactions with signaling molecules. Antioxidants & Redox Signaling **8(11-12,** 2143-2151.

Zhang P., Johnson P. S., Zöllner C., Wang W., Wang Z., Montes A. E., Seidleck B. K., Blaschak C. J., and Surratt C. K. (1999). Mutation of human μ opioid

receptor extracellular "disulfide cysteine" residues alters ligand binding but does not prevent receptor targeting to the cell plasma membrane. Molecular Brain Research **72(2),** 195-204

Zhao D., Ni W, Feng B., Han T., Petrasek M. G., Ma H. (2003) Members of the *Arabidopsis-SKP1-like* gene family exhibit a variety of expression patterns and may play diverse roles in Arabidopsis. Plant Physiol **133,** 203–217

Zhao D., Ni W., Feng B., Han T., Petrasek M. G., and Ma H. (2003). Members of the Arabidopsis-SKP1-like gene family exhibit a variety of expression patterns and may play diverse roles in Arabidopsis. Plant Physiol **133(1),** 203-217.

Supplementary Chapter

S.1 Preparation the FLAG-GCR1, Myc-TRX3, Myc-S42, Myc-TRX4, Myc-S43 and Myc-*zf*-DHHC1 constructs for use in the Co-IP

The FALG-GCR1, Myc-TRX3, Myc-TRX4, Myc-S42, Myc-S43 and Myc-*zf*-DHHC1 constructs (Figure S.1) were prepared for use in the Co-IP.

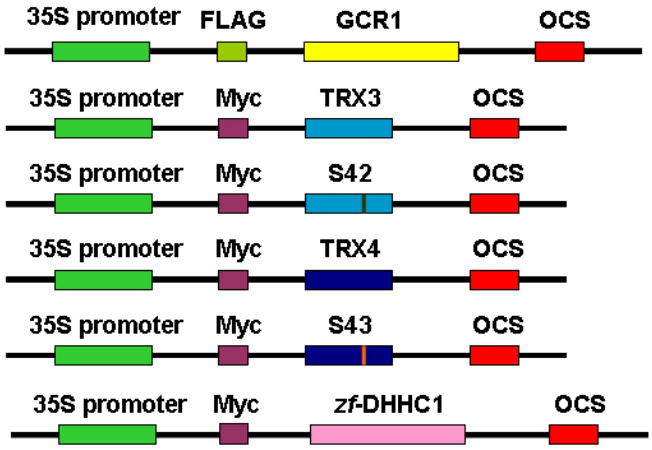


Figure S.1 Schematic representation of the FALG-GCR1, Myc-TRX3, Myc-S42, Myc-TRX4, Myc-S43 and Myc-*zf*-DHHC1 constructs for use in the Co-IP.

The GCR1, TRX3, S42, TRX4, S43 and *zf*-DHHC1 inserts were PCR amplified using the KOD DNA polymerase (section 2.2.2.1). GCR1 was amplified using *GCR1* cDNA as templates, with the GCR1CACCbeg + CterEndHindIII primers pair (Table 2.8). TRX3 and S42 were amplified using pUra-TRX3 and pUra-S42 as templates respectively, with the TRX3CACCbeg+TRX3EndRI primer pair (Table 2.8). TRX4 and S43 were amplified using pUra-TRX4 and pUra-S43 as templates respectively, with the TRX4CACCbeg+TRX4EndRI primer pair (Table 2.8). *zf*-DHHC1 was amplified using pCR2.1-*zf*-DHHC as template with the ZFCACCbeg + ZFEnd primers pair (Table 2.8). The predicted size of each insert is GCR1: 994 bp,

TRX3: 382 bp, S42: 382 bp, TRX4: 385 bp, S43: 385 bp, *zf*-DHHC1: 1027 bp and the products obtained were consistent with this (Figure S.2).

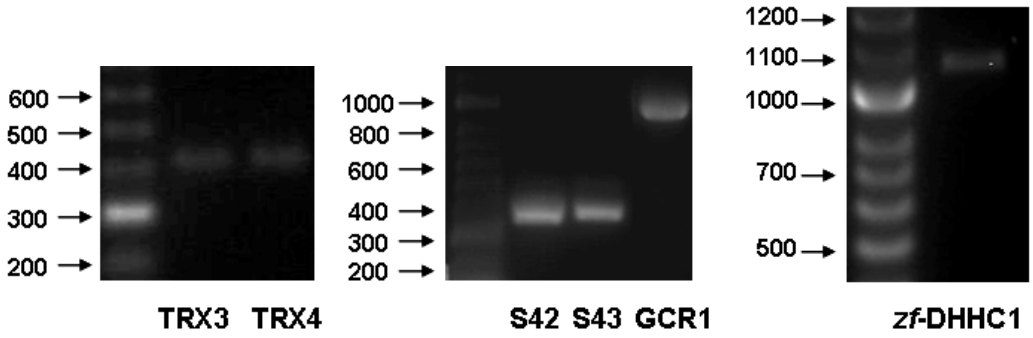


Figure S.2 PCR amplification of the insert. The TRX3, TRX4, S42, S43and GCR1 inserts were PCR amplified. The predicted size of each insert is: TRX3: 382 bp, TRX4: 385 bp, S42: 382 bp, S43: 385 bp, GCR1: 994 bp, *zf*-DHHC1: 1027 bp.

The inserts were cloned into pENTR/D as described in section 2.2.2.6. Recombinants were checked for insert size and orientation by colony PCR with a vector specific M13F as forward primer and an insert-specific primer e.g. CterEndHindIII as reverse primer (data not shown). Positive clones were identified and the plasmids were extracted and sequenced (data not shown). Sequencing results confirmed that all inserts were successfully cloned into pENTR/D, which generated the pENTR/D-GCR1, pENTR/D-TRX3, pENTR/D-S42, pENTR/D-TRX4, pENTR/D-S43 and pENTR/D-*zf*-DHHC1 constructs, which were used as the entry clones in the following LR recombination.

The entry clone and the destination vector, e.g. pEarleyGate202 were used in a LR recombination reaction that was carried out as described in section 2.2.2.7. The reaction mix was then transformed into competent *E.coli* cells and plated onto LB-Kan plates and incubated at 37°C overnight. Recombinants were checked for insert size and orientation by colony PCR with a vector specific 35S as forward primer, and an insert-specific primer e.g. CterEndHindIII as reverse primer (data not shown). Positive clones were identified and the plasmids were extracted and sequenced (data not shown). Sequencing results confirmed that all inserts were successfully cloned in frame to the C terminus of the epitope tag in the destination vectors, which

generated the pEarleyGate202-GCR1 (FLAG-GCR1), pEarleyGate203-TRX3 (Myc-TRX3), pEarleyGate203-S42 (Myc-S42), pEarleyGate203-TRX4 (Myc-TRX4), pEarleyGate203-S43 (Myc-S43) and pEarleyGate203-*zf*-DHHC (Myc-*zf*-DHHC1) constructs.

S.2 Transient expression of Myc-TRX3, Myc-S42, Myc-TRX4 and Myc-S43 in tobacco

The Myc-TRX3, Myc-S42, Myc-TRX4 and Myc-S43 constructs were transformed into *Agrobacterium tumefacien* as described in section 2.2.17. Colony PCR was performed on randomly selected *Agrobacterium tumefacien* colonies as described above to ensure that they each carry the corresponding construct. These colonies were used to inoculate liquid cultures for the infiltration of tobacco plants as described in section 2.2.18.

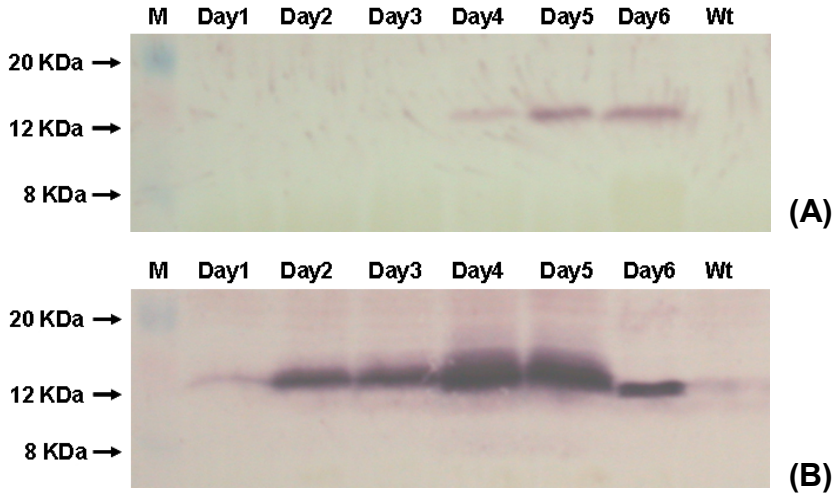


Figure S.3 Western blot checking the expression of Myc-TRX3 in tobacco without (A) and with (B) a silencing suppressor. Leaf discs of infiltrated plants were taken every 24 hours up to day 6. Uninfiltrated wild type tobacco (Wt) was used as a control. The predicted size of the recombinant protein is 14.5 kDa.

The Myc-TRX3 construct was infiltrated into tobacco with or without a silencing suppressor. Leaf discs of infiltrated plants were taken every 24 hours up to day 6, and protein expression was checked by Western blot. As can be seen by comparing figure S.3A and S.3B, the expression level of the

target protein was much higher when co-infiltrated with a silencing suppressor than without it. It indicates that silencing suppressors are effective at inhibiting post-transcriptional gene silencing, and elevating transient expression of the target gene in *Agrobacterium tumefacien* infiltrated tobacco. Therefore, the other constructs were infiltrated into tobacco with a silencing suppressor. The expression of Myc-S42, Myc-TRX4 and Myc-S43 were detected in infiltrated tobacco plants (data not shown).

Appendices

Appendix 1 Vector sequences

A1.1 Multiple cloning site for pMetRas (RRS bait vector)

```
   HindIII SmaI/XmaI   NcoI
5'-AAG CTT CCC GGG ACC ATG GAG CAA AAG CTC ATT TCT GAA GAG GAC
                       M   E   Q   K   L   I   S   E   E   D

TTG AAT TCA AGG GGG ACC ATG – Ras cDNA – TGG ATC CAC TAG – 3'
 L   N   S   R   G   T   M
```

The 17 amino acid Myc sequence is underlined.

Ras cDNA sequence

```
   1  ATGGAATATA AGCTGGTGGT GGTGGGCGCC GGCGGTGTGG GCAAGAGTGC
  51  GCTGACCATC CAGCTGATCC AGAACCATTT TGTGGACGAA TACGACCCCA
 101  CTATAGAGGA TTCCTACCGG AAGCAGGTGG TCATTGATGG GGAGACGTGC
 151  CTGTTGGACA TCCTGGATAC CGCCGGCCTG GAGGAGTACA GCGCCATGCG
 201  GGACCAGTAC ATGCGCACCG GGGAGGGCTT CCTGTGTGTG TTTGCCATCA
 251  ACAACACCAA GTCTTTTGAG GACATCCACC AGTACAGGGA GCAGATCAAA
 301  CGGGTGAAGG ACTCGGATGA CGTGCCCATG GTGCTGGTGG GGAACAAGTG
 351  TGACCTGGCT GCACGCACTG TGGAATCTCG GCAGGCTCAG GACCTCGCCC
 401  GAAGCTACGG CATCCCCTAC ATCGAGACCT CGGCCAAGAC CCGGCAGGGA
 451  GTGGAGGATG CCTTCTACAC GTTGGTGCGT GAGATCCGGC AGCACAAGCT
 501  GCGGAAGCTG AACCCTCCTG ATGAGAGTGG CCCCGGCTGC ATGAGCTGCA
 551  AG
```

A1.2 Multiple cloning site for pUra (RRS prey vector)

```
  HindIII       NcoI
5′-AAGCTTTCTAGACC ATG GGG AGT AGC AAG AGC AAG CCT AAG GAC CCC
                  M   G   S   S   K   S   K   P   K   D   P

                     SmaI  BgI*BamH*
   AGC CAG CGC CGG CCC GGG AGA TCC ACT AGT AAC GGC CGC CAG TGT
    S   Q   R   R   P   G   R   S   T   S   N   G   R   Q   C

         EcoRI                                        XhoI
   GCT GGA ATT CTG CAG ATA TCC ATC ACA CTG GCG CGG GCT CGA G -3′
    A   G   I   L   Q   I   S   I   T   L   A   A   A   R
```

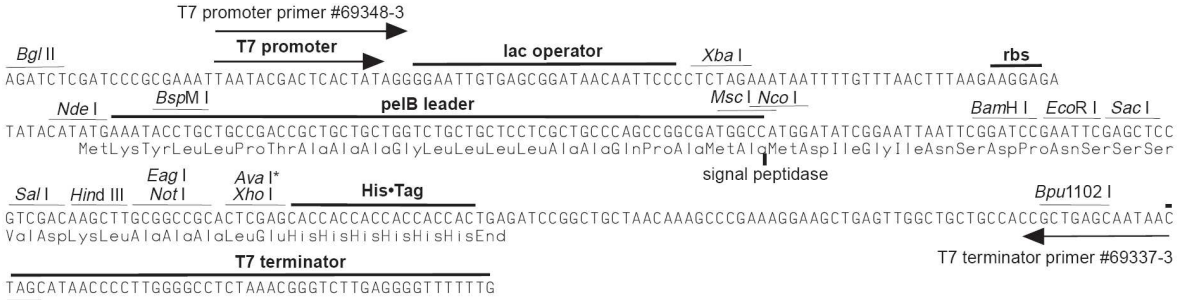
A1.3 Multiple cloning site for pMet (rRRS bait vector)

```
            SmaI/XmaI  PstI                    HindIII
5′-CGA TCC CCC GGG CTG CAG GAA TTC GAT ATC AAG CTT ATC GAT ACC


 SalI    XhoI
GTC GAC CTC GAG TCA TGT - 3′
```

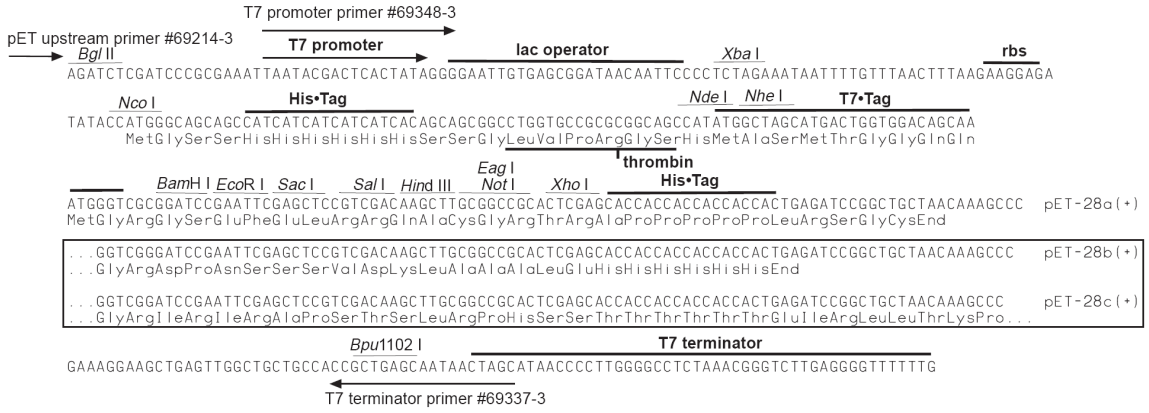
A1.4 Multiple cloning site for pUraRas (rRRS prey vector)

```
     Hind3 SmaI/XmaI   NcoI
5′-AAG CTT CCC GGG ACC ATG GAA TGC ATG – Ras cDNA – TGC AAG


     EcoRI                                        XhoI
   CGA ATT CTG CAG ATA TCC ATC ACA CTG GCG GCC GCT CGA G – 3′
```

A1.5 Cloning and expression region of pET22b(+) (taken from Novagen User Protocol TB038)



A1.6 Cloning and expression region of pET28a(+) (taken from Novagen User Protocol TB074)



Appendix 2 Media composition

Glucose-based medium (pH5.6)	1 Liter
Yeast nitrogen base (without amino acids and NH_4SO_4)	1.7 g
NH_4SO_4	5 g
Glucose	20 g
Bactoagar	25 g
Tryptophan	0.05 g
Lysine	0.05 g
Histidine	0.05 g
Adenine	0.10 g
Leucine (L)*	0.05 g
Uracyl (U)*	0.05 g
Methionine (M)*	0.05 g

* L and M are excluded from Glu-L-M medium; L is excluded from Glu-L+M medium; U is excluded from Glu-U+M medium; L and U are excluded from Glu-L-U+M; L, U and M are excluded from Glu-L-U-M medium.

Galactose-based medium (pH5.6)	1Liter
Yeast nitrogen base (without amino acids and NH_4SO_4)	1.7 g
NH_4SO_4	5 g
Galactose	20 g
Rafinose	20 g
Glycerol	20 g
Bactoagar	25 g
Tryptophan	0.05 g
Lysine	0.05 g
Histidine	0.05 g
Adenine	0.10 g
Leucine (L)*	0.05 g
Uracyl (U)*	0.05 g
Methionine (M)*	0.2 g

* U is excluded from Gal-U+M medium; L and U are excluded for Gal-L-U+4M medium; L, U and M are excluded form the Gal-L-U-M medium.

YPD medium*	1 Liter
Yeast extract	10 g
Peptone	20 g
Glucose	20 g

2x YPD medium	1 Liter
Yeast extract	20 g
Peptone	40 g
Glucose	40 g

LB medium (pH7.0)*	1 Liter
Tryptone	10 g
Yeast extract	5 g
NaCl	5 g

2xYT medium (pH7.0)	1 Liter
Tryptone	16 g
Yeast extract	10 g
NaCl	5 g

SOC medium **	1 Liter
Tryptone	20 g
Yeast extract	5 g
NaCl	0.5 g

NZY^{+} medium (pH7.5)***	1 Liter
NZ amine (casein hydrolysate)	10 g
Yeast extract	5 g
NaCl	5 g

*1.7% agar was added for solid media

** Add the following filter-sterilized supplements prior to use: 10 ml 250 mM solution of KCl and 20 ml of 1 M solution of glucose

*** Add the following filter-sterilized supplements prior to use: 12.5 ml of 1 M $MgCl_2$, 12.5 ml of 1 M $MgSO_4$ and 20 ml of 20% (w/v) glucose.

Appendix 3 Primer sequences

Primer	Sequence* (5' – 3')	T_m (°C)
GCR1bgn	ATG TCG GCG GTT CTC ACA GC	65.9
GCR1begHindIII	CCC AAG CTT ACC ATG TCG GCG GTT CTC AC	79.4
GCR1begHindIII 2	CCC AAG CTT CT ACC ATG TCG GCG GTT CTC AC	80.3
GCR1CACCBeg	CACC ATG TCG GCG GTT CTC ACA GC	77.2
LP1R	GAC AAC AGT ACG ATG CTT GAA AGA GAA TTT TCG AAG	72.1
LP2F	CAT CGT ACT GTT GTC AAG C	51.9
LP3R	AAG CTT TCT GTT CAA CAC CTT TAA CTC	61.5
LP1FXmaI	TCC CCC GGG ACC ATG GCG AAA GAA CTT CG	84.1
LP1RXmaI	TCC CCC GGG CTT GAA AGA GAA TTT TCG	77.5
LP2FXmaI	TCC CCC GGG ACC ATG GCG CAT CGT ACT GTT GTC AA	88.3
LP2RXmaI	TCC CCC GGG TTC CAA ATC CTC CAC ATC AG	81.8
LP3FXmaI	TCC CCC GGG ACC ATG GCG CAA GTG ATA CG	85.6
LP3RXmaI	TCC CCC GGG TCT GTT CAA CAC CTT TAA C	76.1
CterFXmaI	TCC CCC GGG ACC ATG GGT TTC AAC AGC TC	83.2
CterRXmaI	TCC CCC GGG TTG CTG GTC CTC GGT CTT G	84.5
CYCLP1BegEcoRI	CCG GAA TTC ACC ATG GCG TGC TAC TGC CTC	81.9
LP1F-VLC EcoRI	CCG GAA TTC TG ACC ATG GTT CTC TGC TAC TGC	78.2
LP1F-CYC EcoRI	CCG GAA TTC TG ACC ATG GCG TGC TAC TGC CTC	82.6
LP1FEcoRI	CCG GAA TTC ACC ATG GCG AAA GAA CTT CG	78.9
LP1FEcoRInew	CCG GAA TTC TG ACC ATG GCG AAA GAA CTT CG	79.9
LP1REcoRI	CCG GAA TTC CTT GAA AGA GAA TTT TCG	71.0
LP1REcoRInew	CCG GAA TTC C CTT GAA AGA GAA TTT TCG	73.1
LP1End EcoRI	CCG GAA TTC TCA CTT GAA AGA GAA TTT TCG	72.4
LP1GGGrev	ATG ACC GCC ACC CTT GAA AGA GAA TTT TCG	76.9
GGGLP2for	AAG GGT GGC GGT CAT CGT ACT GTT GTC AAG	77.6
LP2BegEcoRI	CCG GAA TTC ACC ATG GCG CAT CGT ACT GTT G	81.9
LP2FEcoRI	CCG GAA TTC TG ACC ATG GCG CAT CGT ACT GTT G	82.7
LP2REcoRI	CCG GAA TTC C TTC CAA ATC CTC CAC ATC AG	77.5
LP2End EcoRI	CCG GAA TTC TCA TTC CAA ATC CTC CAC ATC	76.1
LP2endHindIII	CCC AAG CTT TCA TTC CAA ATC CTC CAC ATC	76.0

Primer	Sequence* (5' – 3')	T_m (°C)
LP3begHindIII	CCC AAG CTT ACC ATG GCG CAA GTG ATA CGG ATG CT	82.8
LP3BegHindIII 2	CCC AAG CTT CT ACC ATG GCG CAA GTG ATA CGG ATG CT	83.4
LP3EndHindIII	CCC AAG CTT TCA TCT GTT CAA CAC CTT TAA C	72.2
CterBegHindIII	CCC AAG CTT CT ACC ATG GGT TTC AAC AGC TCA GTG	79.7
CterendHindIII	CCC AAG CTT TCA TTG CTG GTC CTC GGT C	78.5
TRX3Beg	ATG GCC GCA GAA GGA GAA GTT ATC GCT TGC	78.5
TRX3R1	TCA AGC AGC AGC AAC AAC TGT C	67.7
TRX4Beg	ATG GCG GCA GAA GAG GGT CAA GTG ATT GG	79.3
TRX4R1	TTA CGC AGT TGT AAC ACC AGT ATG	62.6
TRX3BegEcoRI	G GAA TTC G ACC ATG GCC GCA GAA GGA G	79.4
TRX3BegBmRI	GGAA TTC TG GGA TCC ACC ATG GCC GCA GAA GGA GA	85.3
TRX3EndRI	G GAA TTC GA GAG AAA AGC AAA TCA AGC	69.9
TRX4BegEcoRI	G GAA TTC G ACC ATG GCG GCA GAA GAG G	79.4
TRX4BegBmRI	G GAA TTC TG GGA TCC ACC ATG GCGGCAGAAGAGG	84.7
TRX4EndRI	G GAA TTC TGC CTC AAA CTG ATT TAC G	68.1
$TRX3^{Ser39}F$	GCA ACA TGG TCA CCA CCT TGC CG	76.2
$TRX3^{Ser39}R$	CG GCA AGG TGG TGA CCA TGT TGC	76.2
$TRX3^{Ser42}F$	CCA CCT TCA CGT TTC ATT GCA CC	71.3
$TRX3^{Ser42}R$	GG TGC AAT GAA ACG TGA AGG TGG	71.3
$TRX4^{Ser40}F$	GCT TCA TGG TCA CCA CCA TGC CG	76.3
$TRX4^{Ser40}R$	CG GCA TGG TGG TGA CCA TGA AGC	76.3
$TRX4^{Ser43}F$	CCA CCA TCA CGC ATG ATT GCT CC	74.5
$TRX4^{Ser43}R$	GG AGC AAT CAT GCG TGA TGG TGG	74.5
TRX3CACCBeg	CACC ATG GCC GCA GAA GGA GAA G	74.5
TRX4CACCBeg	CACC ATG GCG GCA GAA GAG GGT C	76.4
ZFCACCBeg	CACC ATG GGC GTT TGT TGC	69.7
ZFBegHindIII	AAG CTT ACC ATG GGC GTT TGT TGC C	74.3
ZFEnd	TTA GCA GCA GCG ACA TTT C	61.7
MT F1	GTG TTA ACC ATA TGT GAG GGT G	60.6
MT R1	TTA ACG CAT TTT GTA GTG TTG GG	64.4
CL12A Beg	ATG GCG TCG ACG ACT CTC TCA ATC G	74.8
RPL27aC F1	CCT CTT CGA CAA GTA CCA TCC	62.4

Primer	Sequence* (5' – 3')	T_m (°C)
PRL27aC R1	TTA AGC AGT AAG CAC AAC AGC	59.7
CL12C F1	TCG GTG TCT CCC CAC TCT CCT	69.5
CL12C R1	TTA AGC AAT GGA GAC TTT AGC ACC	64.1
PAG1 F1	TGG CGA AAA TCA TCT ACA AGC	63.8
PAG1 R1	TTA GTC AGC ATC CAT CTC CTC	60.1
BetvI KpnIBeg	GGT ACC ATG GTA GAG GCA GAG GTT GAA G	71.5
BetvI R1	CTA ACC CTC AGA CAA TAG GTA TTG	59.5
AGL42 F1	CCA TGG AGA GGA TCA TCG ACT G	66.3
AGL42 R1	CTA GCA GTT TCT ATT TGG CA	57.2
TRX2 F1	ACA GCT GCA GGG ACC GAA TC	68.8
TRX2 R1	TTA TGC TCT GAG TTT GCT AAC	56.3
ZF F1	GCT CGT GAT CAA TGT CTC CTC	63.8
ZF R2	GTC AAC GTA CAT GAT GAG CG	61.8
ASK1 F1	CTG GCT GCT AAT TAC CTG AAT	59.8
ASK1 R1	TCA TTC AAA AGC CCA TTG GT	64.1
AHB2 F1	TGG TAG TGG CTG ACA CAA	58.7
AHB2 R1	TTA TGA CTC TTC TTG TTT CAT CTC G	62.3
Thionin Beg	ATG AAA GGA AGA ATT TTG AT	54.3
Thionin R1	TTA CAA CAG TTT AGG CGG C	60.3
QPRTase F1	CTG GAT AAT ATG GTT GTG CC	59.3
QPRTase R1	CTA TGC TCG TTT GGT CCT TC	61.2
LRR RLK F1	GCT ATC AAC TGT CTT TGC GA	60.6
LRR RLK R2	GGT AGT CCA TGA AAC TAC ATC T	56.2
SRO5 F1	TGA TCA AAG CAT TGT CCA AGT	60.1
SRO5 R1	TTA GTG TTG TAC TTT ATG TCC ACA	57.7
ZAC F1	GCT GAA ATC GCA CGA CAA TG	66.6
ZAC R1	TTA TTG CTC AAG AGG TAG CCA	61.3
40SRPS2 Beg	ATG GCG GAA AGA GGA GGA GAA	68.1
40SRPS2 R1	TTA AGC TTG GTC TTC ACC CTC	62.1
ExPro F1	GTA CCA CAC ACT TCC GTT GG	62.9
ExPro R1	CTA AGC GAC ATG AGA AAC CAC	61.2
P450 F1	GTT CAA GAG TTT GAG TGG AGC GCT	67.2

Primer	Sequence* (5' – 3')	T_m (°C)
P450 ER1	TTA GAC CCT TGG CTT AAC CAT	61.2
PDF2.2 Beg	ATG AAG CTC TCT ATG CGT TTG	60.9
PDF2.2 ER1	TCA GCA ATG TCT GGT GC	59.4
C2dcp F1	CAC ACA AGA CTT GAT CCT TAG AC	58.3
C2dcp R1	AGA GCC TCC ACC TTT G	55.9
MLPR F1	GGT GAT CTG ATG AAT GAG TAC AAG	61.3
MLPR ER1	CTA TTC CTC GGC CAA GAG ATG TTC G	71.1
ADPGluP F1	GAG GGA GTC TAC AAT TCG AAA C	60.5
ADPGluP ER1	TTA CAA AAT GAT CTC GTC TTG AAC AC	64.0
Glyrich F1	GTG AAG GAG GTG GCT ATG GAG GA	68.8
Glyrich R1	TGA CCA CCA CCA CCA CCG TGA CCT	77.5
FLAGIC	CCTGTCATCG TCGTCCTTGT AGTC	67.8
T7	AAT ACG ACT CAC TAT AGG G	50.4
pUraR	CTT CCT TTT CGG TTA GAG CG	62.6
pMetF	AGC GTC TGT TAG AAA GGA AGT	58.6
pMetRI	CTT CCT TTT CGG TTA GAG C	58.1

*The restriction enzyme sequences are underlined. The sequence encoding the ATG start codon is double underlined. Additional sequence GCG (wavy lined) is added after the start codon in certain i1 and i2 primer sequences to ensure high transcription levels.